COLLEGE PHYSICS

THE MACMILLAN COMPANY
NEW YORK • BOSTON • CHICAGO
SAN FRANCISCO

MACMILLAN & CO., LIMITED
LONDON • BOMBAY • CALCUTTA
MELBOURNE

THE MACMILLAN CO. OF CANADA, LTD.
TORONTO

COLLEGE PHYSICS

BY

JOHN OREN REED, PH.D.

PROFESSOR OF PHYSICS IN THE UNIVERSITY OF MICHIGAN AND DEAN OF THE DEPARTMENT OF LITERATURE, SCIENCE AND THE ARTS

AND

KARL EUGEN GUTHE, PH.D.

PROFESSOR OF PHYSICS IN THE UNIVERSITY OF MICHIGAN

New York

THE MACMILLAN COMPANY

1924

S

Set up and electrotyped. Published July, 1911.

Norwood Press
J. S. Cushing Co. — Berwick & Smith Co.
Norwood, Mass., U.S.A.

PREFACE

In the preparation of this text the authors have kept constantly in mind three distinct purposes, which seem to them to be of paramount importance in any textbook: (*a*) to present the fundamental facts of physics in clear, concise and teachable form; (*b*) to relate these fundamental facts to the basic laws and to the theories of physics in such way as to render plain the historical growth of the science; and (*c*) to put the student in direct touch with firsthand information concerning the epoch-making discoveries of the past upon which the growth of the science has been based, as well as to afford an intimation of the marvelous progress of the present.

In pursuance of the first of these purposes, that arrangement of topics has been chosen which, in the experience of the authors, has been found to lend itself most readily to a simple and natural presentation of the subject as a whole. Owing to the more obvious relations existing between them, the subject of heat is made to follow immediately after the distinctly material phenomena of mechanics and sound; electricity precedes light, and the subject of radiation, usually found under the different chapters of heat, electricity and light, is treated separately after these subjects have been presented. It has also been thought best, even at the sacrifice of historical consistency, to begin the subject of electricity with current electricity, in order to secure the advantage of the greater familiarity of the student with the phenomena of applied electricity.

Again, it has been deemed wise to preface the treatment of the various subjects with a brief but explicit statement of the different theories which have contributed to the progress of the science. In some cases attention has merely been called to the newer theories, where it has not been considered

advisable to insert an extended treatment in a textbook. It is hoped that this theoretical background may serve to bring out in sharper relief the established laws of physics which are true regardless of any assumption or hypothesis by means of which their explanation may have been attempted.

In accordance with their third purpose, the authors have attempted to put the student in touch with the history of the science, through numerous references to original papers. It is hoped that such references may serve to add to the interest in the study as well as to provoke a spirit of inquiry into the methods employed and the validity of the conclusions reached.

Special effort has also been made to bring within the comprehension of the average college student the results of modern theories and recent investigations. To this end the electron theory, radioactivity and radiation have been given somewhat more than usual prominence.

In cases where it may be found necessary to shorten the course, the paragraphs marked with a star may be omitted. Throughout the text there will be found references to laboratory experiments, as described in Reed and Guthe's "Manual of Physical Measurements," 3d edition, George Wahr, Ann Arbor.

The authors desire to express their thanks to Professors L. P. Sieg, of the State University of Iowa, C. W. Greene, of Albion College, and W. W. Beman, of the University of Michigan, for numerous valuable criticisms and suggestions. Thanks are also due to their colleagues in the Department of Physics in the University of Michigan for their cordial interest and helpful suggestions.

JOHN O. REED.
KARL E. GUTHE.

ANN ARBOR,
June, 1911.

CONTENTS

MECHANICS

INTRODUCTORY

CHAPTER I

FUNDAMENTAL PRINCIPLES

MECHANICS OF SOLIDS

CHAPTER II

FORCE AND MOTION

CHAPTER III

Types of Motion

CHAPTER IV

Work and Energy

CHAPTER V

Mechanics of a Rigid Body

CHAPTER VI

Elasticity

MECHANICS OF FLUIDS

CHAPTER VII

Fluids at Rest

CHAPTER VIII

Fluids in Motion

MOLECULAR MECHANICS

CHAPTER IX

Surface Phenomena

CHAPTER X

Solution and Diffusion

SOUND

ORIGIN AND PROPAGATION

CHAPTER XI

Nature of Sound

CHAPTER XII

Velocity of Sound

CHAPTER XIII

Reflection and Superposition of Sound Waves

MUSICAL RELATIONS

CHAPTER XIV

Musical Scales

RESONANCE PHENOMENA

CHAPTER XV

Vibratory Phenomena and Resonance

CHAPTER XVI

Vibration of Air in Pipes and Cavities

HEAT

INTRODUCTION

CHAPTER XVII

Nature of Heat

TEMPERATURE

CHAPTER XVIII

Thermometry

CHAPTER XIX

Expansion

CHAPTER XX

Influence of Temperature upon the Pressure of a Gas

QUANTITY OF HEAT

CHAPTER XXI

Calorimetry

CHAPTER XXII

The Mechanical Theory of Heat

CHAPTER XXIII

Transformation of Heat into Mechanical Energy

CHANGE OF STATE

CHAPTER XXIV

Fusion

CHAPTER XXV

Vaporization

CHAPTER XXVI

Hygrometry

CHAPTER XXVII

Liquefaction of Gases

DISTRIBUTION OF HEAT

CHAPTER XXVIII

Conduction

CHAPTER XXIX

Convection

ELECTRICITY AND MAGNETISM

MAGNETISM

CHAPTER XXX

Action-at-a-distance Theory

CHAPTER XXXI

The Ether-strain Theory

CHAPTER XXXII

Magnetic Field of the Earth

ELECTRODYNAMICS

CHAPTER XXXIII

Fundamental Electrical Units

CHAPTER XXXIV

Ohm's Law and its Applications

CHAPTER XXXV

Electrolysis

CHAPTER XXXVI

Electric Cells

CHAPTER XXXVII

Thermoelectricity

CHAPTER XXXVIII

Application of the Heating Effect of Currents

CHAPTER XXXIX

Electrical Condensers

CHAPTER XL

Electromagnetics

CHAPTER XLI

Electromagnetic Induction

CHAPTER XLII

Dynamo-electric Machines

ELECTROSTATICS

CHAPTER XLIII

FUNDAMENTAL PHENOMENA

CHAPTER XLIV

The Electrostatic Field

CHAPTER XLV

Electrostatic Capacity

THE ELECTRON THEORY

CHAPTER XLVI

Electrolytic Conduction

CHAPTER XLVII

Conduction through Gases

CHAPTER XLVIII

Radioactivity

LIGHT

INTRODUCTION

CHAPTER XLIX

Fundamental Phenomena

GEOMETRICAL OPTICS

CHAPTER L

Phenomena of Reflection

CHAPTER LI

Phenomena of Refraction

CHAPTER LII

Prisms and Lenses

CHAPTER LIII

Dispersion

CHAPTER LIV

Optical Instruments

PHYSICAL OPTICS

CHAPTER LV

Velocity of Light

CHAPTER LVI

Interference

CHAPTER LVII

Diffraction

CHAPTER LVIII

Resolving Power of Optical Instruments

CHAPTER LIX

Polarization

CHAPTER LX

Experimental Demonstrations

RADIATION

CHAPTER LXI

Fundamental Laws of Radiation

CHAPTER LXII

Radiation and Temperature

CHAPTER LXIII

Color

CHAPTER LXIV

Electric Waves

LIST OF TABLES

COLLEGE PHYSICS

MECHANICS

INTRODUCTORY

CHAPTER I

FUNDAMENTAL PRINCIPLES

1. Science and Natural Law. Physical Science is concerned with the discovery, investigation, description and explanation of phenomena in the inorganic, or inanimate, world.

The natural tendency of the human mind is to try to arrange the facts of daily observation according to some rational plan; to subject them to some general rule; in short, to explain them. A new fact is considered as explained when it has been shown to be in accord with previous knowledge and to conform to some more comprehensive statement of relationship. Thus daily experience shows that all bodies, such as wood, stone, lead, water, etc., if unsupported, fall to the ground, or if supported, they press upon the support; in other words, they have weight. Torricelli recognized as the cause of the pressure of the air the fact, already known to Galileo, that even gases have weight, and he showed that the ocean of air presses upon the earth's surface because of its weight. In this way the phenomenon of atmospheric pressure was brought into harmony with the facts of previous knowledge, and with the general proposition that all bodies have weight.

Such a proposition is called a law of nature. It has been found to be true in all cases observed; and while it may here-

after be included in some more general proposition, it can never be shown to be false. Science has been defined as "a body of generalizations so irrefragably true, that while they may be subsequently included in some larger generalization, they can never be overthrown." The law of weight has since been included in Newton's law of gravitation, but it has lost none of its truth thereby.

Physical Science embraces the related branches, Physics and Chemistry. The boundaries of these sciences are separated by no sharp line of demarcation, but overlap in many cases, and many laws are common to both. Their methods of attack are daily becoming more similar as their intimate relation is better understood. Their ultimate problem is the investigation of phenomena and the enunciation of laws pertaining to the constitution of matter and its relation to energy.

2. Matter. Matter may be defined as that which we can perceive by our sense of touch. A mass is a definite quantity of matter. Chemistry is occupied with the investigation of changes in the composition of matter. Its fundamental proposition is that of the "Conservation of mass." In accordance with this principle, it is asserted that the quantity of matter in the universe is constant, and that by no human agency can matter be created or destroyed.

Physics is concerned with matter only in so far as it serves as a *carrier of energy.* The fundamental proposition of Physics is the "Conservation of energy." This proposition asserts that the quantity of energy in the universe is constant; that energy, like matter, is indestructible; and that although it may be transformed and transferred in an endless round of changes, no energy is ever lost—the amount of energy remains the same. This does not mean that all the energy is available, or that it will remain so. Much of the energy at our disposal is wasted, in that it escapes in the form of uniformly diffused heat and is thereby rendered unavailable. It has not, however, been destroyed.

3. Inertia. Of the various properties of matter, such as extension, impenetrability, divisibility, porosity, compressibility,

elasticity, weight and inertia, perhaps the most characteristic is that of inertia. Inertia of a body is its persistence in its condition of rest or uniform rectilinear motion. Matter is powerless of itself either to move or to stop moving if once set in motion; moreover, it resists any attempt to move it if at rest, or to stop it if in motion.

Illustrations of inertia are seen in the hammering of the water in a water pipe on suddenly closing the faucet, in the action of the hydraulic ram or of the fly wheel of an engine. Familiar examples are also found in the stamping of snow from the feet, in the beating of dust from a carpet, in the motion of a bicycle rider when his wheel strikes a stone or in the case of a person who steps from a rapidly moving car while facing to the rear. More remarkable illustrations of inertia are seen in the action of dynamite when exploded upon the surface of a rock— the inertia of the air being sufficient to cause the rock to be pulverized by the sudden pressure; in the method of supplying locomotives with water while running at full speed, and in milling machinery in which rapidly revolving steel bars beat the grain to powder.

4. Fundamental Units and Measurements. The measurement of any concrete or physical quantity consists in comparing it with some quantity of the same kind assumed as the standard or unit. Its magnitude or measure is then stated in terms of that unit, and consists of two parts: a *numerical* part, and the part which *names the unit* with which it has been compared. Both these parts are needed to give an exact idea of the quantity in question. Thus we may give the length of a table as 4.57 *meters*, or as 4.57 *feet*, but no idea of the length of the table is possible until the *unit of length* is stated.

The fundamental concepts of Physics are those of space, mass and time, and most physical quantities may be expressed in terms of these. For this reason the units of *length*, *mass* and *time* are called *fundamental units*, and all other units expressible in terms of these are called *derived units*. Such a system of units is called an absolute system. In the system in common use among scientific men the unit of length is the *centimeter*,

the unit of mass is the *gram* and the unit of time is the *second.* This is usually called the C. G. S. system. The corresponding units in the English system are the *foot*, the *pound* and the *second.*

The *centimeter* is the one hundredth part of the standard meter. The standard meter is represented by the distance, at the temperature of melting ice, between two marks on a certain bar of platinum-iridium, known as the international meter, kept at the International Bureau of Weights and Measures, near Paris. Two copies of this meter, known as the "national prototypes," are kept at the Bureau of Standards in Washington.

The meter was originally intended to be the one ten-millionth part of an earth-quadrant from equator to pole. Subsequent measurements have shown this distance to be 10,000,880 meters. The term *meter*, therefore, refers to the bar of metal and has no relation to the shape or size of the earth. By an act of Congress in 1866, the yard is defined as $\frac{3600}{3937}$ of a meter, hence the relation between the centimeter and the inch is very nearly

$$1 \text{ in} = 2.54 \text{ cm}$$

The *gram* is the one thousandth part of a mass of metal called a kilogram. The international kilogram is also kept at the International Bureau of Weights and Measures, near Paris. Two prototypes of this kilogram are kept at the Bureau of Standards in Washington.

It was intended that the gram should represent the mass of one cubic centimeter of distilled water at its temperature of maximum density [4° C]. Although more exact determinations have shown this relation to be slightly in error, yet for all practical purposes we may regard the mass of one cubic centimeter of distilled water at 4° C as equal to one gram. By the same Act of Congress in 1866, the relation between the kilogram and the pound was declared to be

$$1 \text{ kilo} = 2.2046 \text{ lb.}$$

The *second* is the unit of time employed in scientific measurements and may be defined as the $\frac{1}{86400}$ part of a mean solar day, where a mean solar day denotes the average time between the successive passages of the sun across the meridian, taken throughout the year.

It remains to be noted that while the units of mass, length and time are called fundamental units, there are employed other units not directly reducible to these, though connected with them by certain constants which have been determined by experiment. Such units are the units of temperature, of heat and of luminous intensity.

5. Dimensional Formulae and Derived Units. It is frequently of advantage to express physical quantities in general terms, in order to show more clearly their relation to each other. In such cases we write the symbols $[M]$, $[L]$ and $[T]$, with the proper exponents, to show how the fundamental quantities of mass, length and time enter into the derived units in question. Such a formula is called a *dimensional* formula. Since the dimensional formula is concerned with the *nature* of the quantity, rather than with its magnitude, numerical coefficients do not appear. A few examples of derived units will illustrate.

Area. The area of any plane figure is proportional to the product of two of its linear dimensions; hence the dimensional formula for *any area* is $[L^2]$; i.e. *the square of a length.* Unit area is 1 cm^2.

Volume. Since the volume of any solid is proportional to the product of three of its linear dimensions, the general formula for a volume becomes $[L^3]$; that is, *the cube of a length.* Unit volume is 1 cm^3 or 1 cc.

Density. The density of a body is, by definition, *the mass per unit volume.* It is found by dividing the mass of a body by its volume; hence, the formula for density is

$$\left[\frac{M}{L^3}\right] \text{ or } [ML^{-3}]$$

Unit density is a density of 1 g per cm^3.

Specific Volume. The specific volume of a substance is defined as *the volume per unit mass of the substance.* It is, consequently, the reciprocal of the density of the substance, and its dimensional formula is, therefore,

$$\left[\frac{L^3}{M}\right] \text{ or } [L^3M^{-1}]$$

Unit specific volume is, accordingly, a volume of 1 cm^3 per gram.

6. Dimensional Formulae. Time Relations. *Velocity.* Velocity is *the time rate of motion.* If a body move over a space of s cm in t sec, then the time rate of motion, or average velocity, is given by the equation

$$v = \frac{s}{t} \tag{1}$$

Even if the velocity vary from instant to instant, yet its value at any given instant is perfectly definite, and is obtained by dividing smaller and smaller spaces ds by the correspondingly small times dt, needed to traverse these spaces. The limiting value of this ratio, as s and t grow smaller and smaller,

$$v = \frac{ds}{dt} \tag{2}$$

is, then, the time rate of motion for that instant of time.

Velocity includes the additional idea of motion in a definite direction. Thus, a velocity is stated as 10 cm per second, from north to south, or as 25 $\frac{\text{cm}}{\text{sec}}$ northeast. In cases where direction is either of no importance, or cannot be stated, this time rate of motion is termed *speed.* In this text, speed will be symbolized by $\bar{v}$. Since velocity is stated in units of length per unit of time, the dimensional formula becomes $[LT^{-1}]$. Unit velocity is a velocity of 1 cm per second.

Acceleration. Acceleration is the *time rate of change of velocity.* When the motion of a body is not uniform, the velocity is

no longer constant, but changes every instant. This change in velocity may be a change either in magnitude or in direction. If this change in velocity be uniform, then the rate of change is constant, and is found by dividing the difference between the final and initial velocities, v and v_0, by the time t, during which the change in velocity occurs, or

$$a = \frac{v - v_0}{t} \tag{3}$$

Even if the acceleration be not constant, equation (3) gives the *average acceleration during the time t.* At any instant this time rate of change of velocity is given by the limiting value of this ratio, or

$$a = \frac{dv}{dt} \tag{4}$$

The dimensional formula for acceleration is

$$\left[\frac{LT^{-1}}{T}\right] = [LT^{-2}]$$

Unit acceleration is an acceleration of 1 cm per second per second. The acceleration g, due to gravity, is 980 cm per second per second.

Momentum. Momentum is the quantity of motion possessed by a body, and is measured by the product of its mass and its velocity. Momentum is symbolized by mv, and its dimensional formula is $[MLT^{-1}]$. Unit momentum is possessed by unit mass moving with unit velocity.

Force. A body has no power to change its motion of itself. Any change in the motion of a body, either in magnitude or direction, must be due to some action upon the body, which we term a force. Force may be defined as that which tends to change the motion of a body, and is measured by the *time rate of change of momentum.* Its equation is

$$F = \frac{mv - mv_0}{t} = ma \tag{5}$$

The dimensions of force are $[M\,L\,T^{-2}]$. The unit of force

is the *dyne*. It is that force which will give *unit mass unit acceleration*.

7. Trigonometrical Formulae. The following trigonometrical relations will find frequent application. In any right triangle ABC (Fig. 1), right angled at C, we have by definition

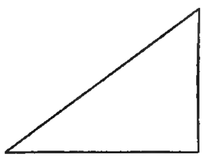

A B C

FIG. 1.

$$\sin A = \frac{\textit{side opposite}}{\textit{hypotenuse}} = \frac{BC}{AB} \tag{6}$$

$$\cos A = \frac{\textit{side adjacent}}{\textit{hypotenuse}} = \frac{AC}{AB} \tag{7}$$

$$\tan A = \frac{\textit{side opposite}}{\textit{side adjacent}} = \frac{BC}{AC} = \frac{\sin A}{\cos A} \tag{8}$$

Also

$$\sin A = \cos (90^\circ - A) = \cos B \tag{9}$$

$$\sin^2 A + \cos^2 A = 1 \tag{10}$$

$$\sin(A + B) = \sin A \cos B + \cos A \sin B \tag{11}$$

$$\sin 2A = 2 \sin A \cos A \tag{12}$$

The following table of values of the sine, cosine and tangent of the various angles should be memorized.

TABLE I

θ	0°	30°	45°	60°	90°	180°
sine	0	$\frac{1}{2}$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}\sqrt{3}$	1	0
cosine	1	$\frac{1}{2}\sqrt{3}$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}$	0	-1
tangent	0	$\frac{1}{3}\sqrt{3}$	1	$\sqrt{3}$	∞	0

8. Circular Measure of Angles. About the vertex C of the angle θ (Fig. 2), describe a circle with radius r, and denote the subtending arc, AB, by s. Then in circular measure the angle θ is defined by the equation

$$\theta = \frac{s}{r} \tag{13}$$

If θ be taken as unity, then $s = r$. This unit angle is called *a radian*. A radian is that angle whose subtending arc is equal to the radius. 1 radian $= 57^\circ.2958 = 3437'.75 = 206265''$.

Again, from the point A drop a perpendicular to CB, and from B erect a perpendicular to CB to meet CA produced. Then, in the two right triangles ACD, and ECB, we have

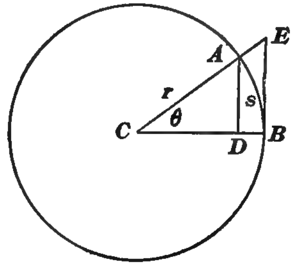


FIG. 2.

$$\sin \theta = \frac{DA}{CA} \tag{14}$$

and

$$\tan \theta = \frac{BE}{CB} \tag{15}$$

Since CA and CB are radii of the same circle, we may write

$$\sin \theta = \frac{DA}{CB},\ \theta = \frac{BA}{CB},\ \tan \theta = \frac{BE}{CB}$$

or the angle θ lies between its sine and its tangent in value; that is,

$$\sin \theta < \theta < \tan \theta \tag{16}$$

As θ approaches 0, these values approach each other, and for small angles, we may write without serious error,

$$\sin \theta = \theta = \tan \theta \tag{17}$$

This approximation[1] is frequently employed in physical formulae. The formulae are usually assumed as correct for values of θ under three degrees.

9. Curvature. The direction of motion of a body at any point in a curved path is the tangent to the curve at that point. Since the direction in which the body is moving constantly changes, it is of importance to know the rate at which the

[1] *For experimental verification of this formula, see, Manual. Exercise 7.*

direction of motion changes. *Curvature is defined as the space rate of change in direction.*

In the case of a point describing the circular arc PQ (Fig. 3), the change in direction is the angle included between the tangents at these points. This angle is in turn equal to the angle θ, included between the radii OP and OQ. But by equation (13)

$$\theta = \frac{s}{r}$$

hence the curvature is

$$\frac{\theta}{s} = \frac{1}{r} \tag{18}$$

or the *curvature* at any point on a curve is the *reciprocal* of the *radius* of the *circle* which most nearly coincides with the curve at that point. In the circle the curvature is constant; it is variable in the ellipse or the parabola, and zero in a straight line.

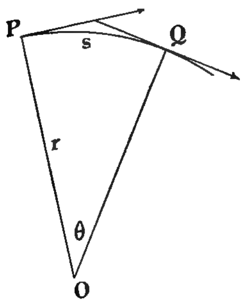


FIG. 3.

10. Vectors and Scalars. Certain physical quantities, such as displacements, velocities, accelerations, forces, etc., are of such a nature that they can be completely described only by giving their *magnitude, direction* and *sense.* Thus consider two elevator cars attached to the opposite ends of the same cable. While their speeds are exactly equal, yet their velocities are very different; although both cars move in the *same vertical direction*, the velocity of the first may be 2 m per second *upward*, while the velocity of the second is 2 m per second *downward.* The *sense* of the velocities is given by the two terms, *upward* and *downward.* If the first velocity be taken as positive, the second must be considered as negative. Such quantities are called *directed quantities* or *vector quantities. A vector quantity is a quantity which may be completely represented by a straight line of definite length, traced in a specific sense.* Its magnitude is given by the length of the line, its direction by the direction of the line and its sense by an arrow-head showing the sense in which the line was traced.

Other physical magnitudes, such as mass, density, energy, power, electrical resistance, etc., have no relation to direction, and are completely described by stating their magnitude. Such quantities are called *scalar quantities.*

11. Projection upon Rectangular Axes. If we have a pair of rectangular axes XX', YY' (Fig. 4), then any vector may be resolved into two components parallel to the two given axes. Thus let OC represent such a vector making an angle θ with the axis of x. Then OA, the component of OC parallel to the axis of x, is defined by the equation

$$OA = OC \cos\theta \qquad (19)$$

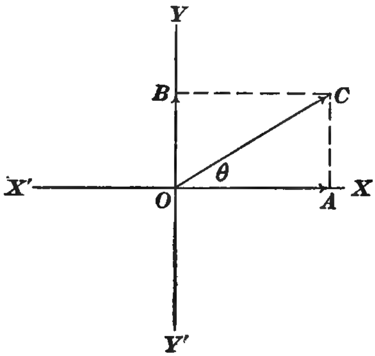


FIG. 4.

Also

$$OB = OC \cos(90° - \theta)$$
$$= OC \sin\theta \qquad (20)$$

These two components, OA and OB, are called the *projections of OC upon the axes of x and y* respectively. The angle θ measures the difference in direction between the line OC and the x-axis and is called the direction angle. *To project any line upon any other, multiply the line in question by the cosine of the direction angle.*

12. Addition and Subtraction of Vectors. Given two vectors p and q, represented in magnitude, direction and sense by the straight lines AB and BC (Fig. 5), as indicated. These may be considered as two displacements, the one from A to B, to which is added another from B to C. Then it is easily seen that these two dis-

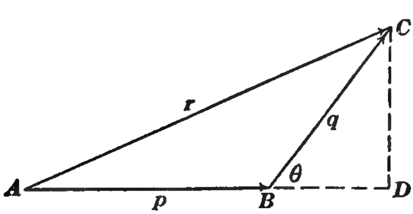


FIG. 5.

placements are equivalent to a single displacement from A to C. Hence AC is defined as the sum r, of the two vectors p and q. This sum may be calculated as follows: From the point C drop a perpendicular upon AB produced, and let the angle θ be the difference in direction between the two vectors. Then

$$\overline{AC}^2 = \overline{AD}^2 + \overline{DC}^2 \tag{21}$$

or

$$r^2 = [p + q \cos\theta]^2 + q^2 \sin^2\theta \tag{22}$$

whence

$$r^2 = p^2 + q^2 + 2pq\cos\theta \tag{23}$$

It is to be noted that this formula covers all possible cases arising from different values of θ. Thus, when $\theta = 0°$, $r = p + q$; for $\theta = 90°$, $r^2 = p^2 + q^2$; $\theta = 180°$, $r = p - q$.

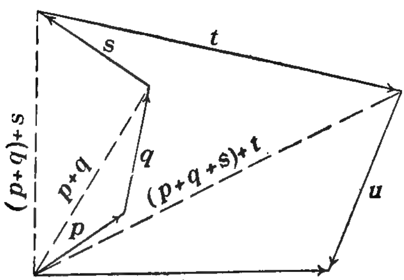


FIG. 6.

The solution may be effected graphically by laying off and connecting in order the lines AB and BC, representing the vectors p and q; then the line AC represents the vector sum of p and q, where the points A, B and C may be any points whatever.

13. Summation of any Number of Vectors. The foregoing graphical method may be extended to the case involving any number of vectors. Thus the resultant r (Fig. 6) of the vectors p, q, s, t and u is represented in magnitude, direction and sense by the line drawn from the beginning of the first to the end of the last, the vectors being added in any order whatever. Thus it is clear that the sum of

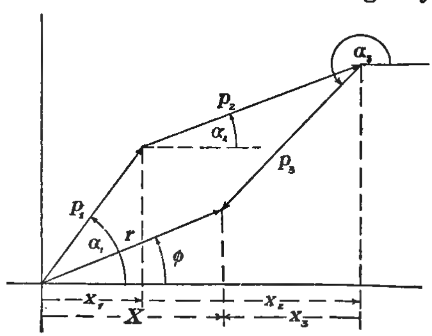


FIG. 7.

any number of vectors is independent of the order in which they are added.

The numerical value of the resultant of any number of vectors is readily calculated by projecting the vectors upon the axes of x and y, adding the various components upon each axis, and then combining the x and y components. Thus, if (Fig. 7) $p_1, p_2, \cdots p_n$, be the various vectors, making angles α_1, $\alpha_2, \cdots \alpha_n$, with the axis of x, then the sum of the x components is

$$x_1 + x_2 + \cdots + x_n = p_1 \cos \alpha_1 + p_2 \cos \alpha_2 + \cdots + p_n \cos \alpha_n \quad (24)$$

or
$$X = \Sigma p \cos \alpha \quad (25)$$

and similarly
$$Y = \Sigma p \sin \alpha \quad (26)$$

whence
$$r^2 = X^2 + Y^2 \quad (27)$$

and the angle ϕ between the resultant and the axis of x is given by the equation

$$\tan \phi = \frac{Y}{X} \quad (28)$$

MECHANICS OF SOLIDS

CHAPTER II

FORCE AND MOTION

14. Force. Our earliest ideas of force are derived from our sense of the muscular exertion needed to produce change in the motion of bodies about us. If we throw a stone into the air, roll a heavy truck along a smooth platform or move a log floating in the water, we are, in each case, conscious of a certain muscular effort needed to put these bodies in motion. Our experience also teaches us that this effort is greater in the case of large bodies than of small ones of the same material, and also that it is more difficult to produce rapid motion in any case than to move the body slowly.

Again, we soon learn that muscular effort is needed to *stop a body* when it is in motion. We learn to consider *rapid motion*, or the *motion of large bodies*, as resulting from *great muscular effort*. In short, we early learn that muscular effort is needed to *change the motion of a body*.

Whenever a body has its condition of motion changed, either in magnitude or in direction, this change is attributed to the action of something which we term *a force*. A force is thus considered as an action upon a mass, and is measured by the *product of the mass and the acceleration conferred upon the mass*, or

$$F = Ma \tag{29}$$

The C. G. S. unit of force is the *dyne*. A dyne is that force which will give a gram mass an acceleration of 1 cm per sec-

ond per second, or *a dyne is that force which will give to unit mass unit acceleration.*

By an easy association of ideas, we come to attribute the motion of bodies to forces other than those directly due to our own muscular efforts. Thus, we say that the branches of the trees are tossed about by the force of the wind, that the motion of the falling body is due to the force of gravity, that the body attached to a spiral spring is kept from falling by the elastic force of the spring or that the ball is driven from the barrel of an air gun by the elastic force of the compressed air.

15. Pressure, Stress, Tension. When a force is distributed over an area, we are accustomed to specify the force exerted upon each unit of area. *Pressure denotes the force per unit area*, or

$$P=\frac{F}{A} \tag{30}$$

The dimensions of a pressure are, accordingly,

$$\left[\frac{MLT^{-2}}{L^2}\right]=[ML^{-1}T^{-2}]$$

The pressure on the piston head of a steam engine is expressed in $\frac{\text{dynes}}{\text{cm}^2}$, or in $\frac{\text{lb}}{\text{in}^2}$

At any point in the interior of a medium subjected to an external force, there exists a system of resisting forces which is termed a *stress.* A stress is likewise measured in terms of *force per unit area.* The dimensions of a stress are the same as those of a pressure.

It is frequently of advantage to discriminate between external and internal pressure. Thus a gas when compressed by an external force reacts against the compressing piston with a pressure equal to the external pressure. If the external forces acting upon a medium be directed toward each other, the medium is said to be under pressure, while if the forces be directed away from each other, it is said to be under tension.

Tension, like pressure, is usually measured in force per unit

area, although the term *tension* is also used in a different sense in the case of *surface tension*, namely, as force per unit length. (See Art. 89.)

In all cases of stress, pressure or tension, the total force exerted upon any area is at once obtained by multiplying the value of the stress, pressure or tension by the area involved, as equation (30) clearly shows.

16. Impulse, Weight, Gravitation and Inertia. *Impulse.* When a force acts for but a short time, as in the case of a blow or a collision, the effect is called an *impulse.* An impulse is measured by the product of the force and the time during which the force acts, or

$$Impulse = Ft \tag{31}$$

The element of time is essential to the consideration of the effect of any force, since nothing short of an infinite force could produce an effect in zero time.

Weight. Again, every force is to be considered as resulting from the mutual action of two bodies. Under this aspect there exists *a stress* in the medium *between the two bodies.* Thus a mass of 1 kilogram is attracted towards the earth and in turn attracts the earth with a force of 980,000 dynes. This mutual action tends to produce motion in the case of each body. The force by which a body is attracted toward the earth is called its *weight*, and its weight is the product of its mass and the acceleration due to gravity, or

$$W = Mg \tag{32}$$

The force 980,000 dynes is the weight of a kilogram, and is sometimes termed *a kilogram weight.* This force is sometimes used as a unit of force, just as the weight of a pound mass may be used as a unit of force. Such units are termed *gravitational units.*

Since the value of g increases slightly from the equator to the pole, it follows that the weight of a body is not constant at different points upon the earth. The *mass* of the body, however, *is constant.* For this reason the C. G. S. units, being independent of any value of g, are frequently termed *absolute units.*

Gravitation. Weight is used to designate the attraction between the earth and different bodies upon its surface. Gravitation is the most general term used to denote the attraction existing between different material bodies anywhere in the universe. Thus gravitational attraction is a general property of all matter. Prior to the time of Newton (1641–1727) but little was known regarding this subject. From mathematical computation Newton showed that the force of attraction F, between two material particles of masses m_1 and m_2, separated by a distance d, is expressed by the equation

$$F = G \cdot \frac{m_1 m_2}{d^2} \tag{33}$$

In this expression G is called *the constant of universal gravitation* and has been found to have an approximate value of 6.48×10^{-8}. By the application of equation (33) Newton was able to account for the motion of the moon about the earth.

Inertia. Weight is not to be confused with inertia, which causes the resistance which a body offers to being set in motion *in any direction.* The weight of a body may be determined by means of the ordinary spring balance. When the body attached to the spring comes to rest, the force Mg, due to the mutual attraction between the body and the earth, is balanced by the elastic force due to a definite distortion of the spring. This force represents the *weight* of the body.

If, now, the balance be suddenly given an acceleration a, upward, the *inertia* of the body causes an *increased stretch* in the spring, and this effect is measured by Ma. This force, due to inertia, may be increased indefinitely as a is increased. Inertia is the cause of the *kinetic reaction* of a body against a change of motion. It is a constant and characteristic property of the body, proportional to its mass; the weight of the body depends on the acceleration due to gravity.

17. Motion. Displacement is a change of position without regard to time. Motion is a change of position occurring in time. All motion is purely relative. Neither absolute rest nor absolute motion is known in the universe. Motion embodies

the two concepts of space and time, and may be conveniently subdivided in accordance with these fundamental relations.

1. *Space Relations.* From the point of view of space relations motions may be said to be of two kinds: (*a*) motion of translation, (*b*) motion of rotation.

2. *Time Relations.* Under this aspect, motion may be studied in its relation to acceleration, from which we have: (*a*) uniform motion, (*b*) uniformly accelerated motion, (*c*) simple harmonic motion.

Besides these simple relations there exist very many combinations, only a few of which can be noticed here.

Space Relations. (*a*) *Translation.* If we imagine a particle to move in space, its path is a *line*, either straight or curved. Such motion is termed *linear motion*, and the displacement, velocity and acceleration concerned are *linear* in character in each case. Such motion is pure translation. If now an extended rigid body move in such a way that each point in the body traces a *right line*, then the body is said to undergo *translation.* Examples of translation are seen in the up-and-down motion of an elevator, or in the motion of a train of cars on a straight level track.

(*b*) *Rotation.* If, on the other hand, the body move so that each point in the body describes a circle about a certain line, then the body is said to *rotate* and the motion is one of *rotation.* The line about which all the points in the body describe circles is called the *axis of rotation.* Since all these circles are described in the same time, it follows that the radii of these various circles all sweep out angle *at the same rate.* Rotation is therefore angular motion, and the displacement, velocity and acceleration concerned are all *angular* in character.

Examples of rotation are seen in the motion of the fly wheel of an engine, or in the spinning of the wheel of a bicycle when held free from the ground.

In nature these two kinds of motion are rarely found entirely distinct from each other. A stick or a ball, when thrown into the air, undergoes both rotation and translation at the same time. Examples of this sort of motion are seen in the motion of a base-

ball when struck "foul," or in the motion of a carriage wheel as it rolls along the ground.

Time Relations. (*a*) *Uniform motion.* If a body undergo either translation or rotation under circumstances such that the acceleration is constantly zero, we have the condition for uniform motion. The motion may be either uniform linear motion or uniform angular motion.

(*b*) *Uniformly accelerated motion.* If translation or rotation occur under circumstances such that the acceleration has and maintains a constant value, we shall have uniformly accelerated motion, of translation or of rotation, as the case may be.

(*c*) *Simple harmonic motion.* In this type of motion the acceleration is directly proportional to the displacement from the position of rest of the body. The resulting motion is either linear or angular simple harmonic motion, according as the displacement from the position of rest was a linear or an angular displacement. To the study of this form of motion several subsequent articles will be devoted.

Besides the various forms of motion already mentioned, there exist in nature numerous combinations, some of which are extremely complex and much too difficult for treatment in an elementary text.

18. Newton's First Law of Motion. We have seen that a body is powerless to acquire motion of itself, and equally incapable of coming to rest of itself, if in motion. Change of motion, therefore, is always due to the action of a force. The equation

$$F = Ma$$

shows that force does not appear except in connection with mass. There is always a mass involved. Again, the factor a indicates that force is exerted *only while the motion is changing.* A steam engine exerts force in pumping water from a well in that it sets the water in motion; the exploding gunpowder exerts force upon the cannon ball during the time the ball is passing from the breech to the muzzle of the gun. The ball in turn exerts force only while its motion is changing, *i.e.*

while it is smashing the target or piercing the armor of the ship. A ball flying through space and encountering no resistance exerts no force; its motion is unchanged. A locomotive pulling a train with uniform velocity along a level track exerts force sufficient to overcome friction, air pressure, etc., but no more. A body in motion moves until some force stops it. This is all summed up in Newton's first law of motion:

"*Every body continues in its state of rest or uniform motion in a straight line, except in so far as it is compelled to change that state by a force impressed upon it.*"

This law is embodied in the equation

$$F = Ma$$

If a be zero, there is no force; hence there must exist either *rest* or *uniform motion.* Again, a denotes the rate of change of velocity *either* in *magnitude or direction;* hence if a be zero, there must exist either *rest* or *uniform motion in a straight line.*

All apparent exceptions to the action of this law are in reality but proofs of this truth. A stone thrown into the air does not move with uniform velocity in a straight line for a single instant; yet the reason is found in the impressed force of gravitation, which compels it to change that state.

19. Newton's Second Law of Motion. The second law of motion is:

"*Change of motion is proportional to the moving force impressed, and takes place in the direction in which the force acts.*"

It is to be noted that by "change of motion" Newton meant *change in momentum*, as his own explanation of the meaning of the law clearly shows. He explains this law as follows: "If a force generate any motion, a double force will generate a double motion, a triple force a triple motion, whether they be applied simultaneously and at once, or gradually and successively. This motion, if the body were already moving, is either added to the previous motion, if in the same direction, or subtracted from it, if directly opposed, or compounded with it if the two motions are inclined at an angle."

This law is also embodied in the equation

$$F = \frac{Mv - Mv_0}{t} = Ma$$

and gives us a means of measuring either force or mass as the case may be. Thus we judge of the mass of a body by the force necessary to set it in motion. We kick a barrel lying on the ground to see whether it is empty or not. If full, it is started with difficulty; if empty, it moves very readily.

Again, suppose we have small cubical blocks of cork, aluminium, and lead, each mounted upon a little car so as to move readily upon a smooth table. We attach to each car a small spring balance and tie the balance to a rod by which we pull the blocks quickly along the table. The result will be that, while all the blocks have practically the same acceleration, the balances will indicate by their stretch the kinetic reactions, that is, the relative masses, of the various substances. In this case, the *acceleration* being the same for all, the *force* is *directly proportional* to the *mass.*

If, on the other hand, the blocks with their cars be placed upon a smooth table and be struck equal blows, as from a spring hammer, then the accelerations produced in the various blocks will afford a measure of their relative masses. The cork block would move off rapidly, the aluminium more slowly and the lead would be moved least of all. We should conclude that the lead contains the most matter, the aluminium next, and the cork least. In this case the *force* is kept constant, and the *accelerations* vary *inversely* as the *masses.*

20. Newton's Third Law of Motion. The third law of motion states that

"*To every action there is an equal and contrary reaction, or the mutual actions of two bodies are equal and opposite.*"

In explanation of this law Newton adds, "Whatever presses or draws another body is pressed or drawn to the same extent by that body. If one press a stone with the finger, the finger is pressed by the stone. If a horse pull on a stone by a rope, the horse is pulled equally toward the stone; . . . and to the extent that the forward motion of the one is aided the forward motion of the other is impeded."

This law expresses the two-sided nature of every force. Force is always due to the mutual action of two bodies, the action of the one being equaled by the reaction of the other. This amounts to saying that forces always occur in pairs.

Again, no force can be exerted unless there be some *resistance* to overcome. There can be no action unless there be something to act upon which will, in its turn, react. The athlete prefers to jump from a slab of stone; he cannot "rise" from a pile of straw or a heap of cushions. For the same reason it is tiresome to walk in melting snow or loose sand.

If motion ensue as the result of the action of two bodies, then the law expresses the equality of the resultant motions; that is, if a force confer upon the masses m and m', velocities v and v', then the law states that their *momenta are equal*, or

$$mv = m'v' \tag{34}$$

Illustrations of this law are manifold. The explosive force of the powder drives the ball from the cannon; the two are shot apart, moving with velocities inversely as their masses. The recoil or "kick" of a gun is the greater the more nearly the masses of gun and projectile are made equal to each other.

The screw of a ship drives the water *backward* with a velocity as many times *greater* than the forward velocity of the vessel, as the mass of water moved is *less* than that of the vessel. Boats have been propelled by machinery which pumped water in at the bow and expelled it in a small stream under high velocity at the stern. Again, the blades of the propeller of an aeroplane are much longer and rotate at a much higher speed than those of the propeller of a boat, since the volume of air to be displaced is much greater than the corresponding volume of water needed to furnish an equal reaction. The motions of fishes in the water, and of birds in the air, the ascent of skyrockets and the action of rotary lawn sprinklers are all explained in accordance with Newton's third law of motion.

CHAPTER III

TYPES OF MOTION

21. Uniform Motion. The simplest type of motion is uniform motion. In the case of rectilinear motion the linear *velocity* v remains *constant;* that is, the space traversed in unit time and the direction and sense of motion all remain *unchanged.* Equal spaces are described in equal times. The *acceleration* is therefore zero. We may, from these conditions, write down the equation of uniform motion from definition. We have

$$\frac{s}{t} = v = constant \tag{35}$$

whence

$$s = vt \tag{36}$$

Similarly, if a body rotate uniformly about an axis, then a straight line drawn from the axis to any point in the body may be conceived as sweeping out angle about the axis at a uniform rate. If the *period*, or the time needed to describe a complete revolution, be T sec, then the total angle swept out in that time is 2π radians, and the time rate of generating angle is

$$\frac{2\pi}{T} = \omega \tag{37}$$

where ω is called the *angular velocity.* In uniform rotation the angular velocity is a constant. Hence if any angle θ be described in time t, then the angular velocity ω is defined by the equation

$$\frac{\theta}{t} = \omega = constant \tag{38}$$

whence the equation for uniform rotation becomes

$$\theta = \omega t \tag{39}$$

These conditions for uniform motion, whether linear or angular, involve the permanence of any motion once set up, and the absence of any force to maintain it.

22. Uniformly Accelerated Motion. In the case of uniformly accelerated linear motion the body moves with a constantly and uniformly increasing or decreasing velocity. It passes over unequal spaces in equal intervals of time. The velocity is no longer constant, but the *acceleration*, or the rate at which the velocity changes, *is constant*. Hence we may write

$$\frac{v - v_0}{t} = a = constant \tag{40}$$

The force producing uniformly accelerated motion is also a *constant force*, since

$$F = Ma$$

in which both M and a are constant.

From equation (40) we see that v, the velocity at any time t, is

$$v = v_0 + at \tag{41}$$

or the final velocity v is equal to the initial velocity v_0, plus the *change in velocity* acquired in time t.

The average velocity v' during this interval of time is, of course,

$$v' = \frac{v' + v_0}{2} \tag{42}$$

where v' denotes the constant velocity at which the body would have described the same space in time t. The expression for the space described is readily found from equation (36); thus we have

$$s = v't = \frac{v + v_0}{2} \cdot t \tag{43}$$

or, replacing v by its value from equation (41), we have

$$s = v_0 t + \frac{at^2}{2} \tag{44}$$

Finally, combining equations (40) and (43) to eliminate t, we have

$$v^2 = v_0^2 + 2\,as \tag{45}$$

It is further to be noted that the acceleration may be either positive or negative; in the latter case the acceleration and the motion are oppositely directed, and the acceleration becomes a retardation. In their general form the equations of uniformly accelerated motion become

$$\begin{aligned} v &= v_0 \pm at \\ s &= v_0 t \pm \frac{at^2}{2} \\ v^2 &= v_0^2 \pm 2\,as \end{aligned} \tag{46}$$

If the minus sign is used, s is the distance from the zero position to that at time t, taken in the direction of v_0.

23. Freely Falling Bodies. In the case of a freely falling body equations (46) become, on substituting for the general acceleration a, the acceleration due to gravity ($g = 980$ cm per second per second),

$$\begin{aligned} v &= v_0 \pm gt \\ s &= v_0 t \pm \frac{gt^2}{2} \\ v^2 &= v_0^2 \pm 2\,gs \end{aligned} \tag{47}$$

In case the body start from rest, v_0 is zero, hence

$$\begin{aligned} v &= gt \\ s &= \frac{gt^2}{2} \\ v^2 &= 2\,gs \end{aligned} \tag{48}$$

It is to be noted that in case of a body starting from rest the velocity is proportional to the time, the space described is proportional to the square of the time, and the space described in the consecutive seconds varies as the odd numbers 1, 3, 5, etc.

If a body be thrown vertically upward with an initial veloc-

ity v_0, the motion and the acceleration are oppositely directed and the lower sign is to be used in equations (47). The time of ascent and the height to which a body thrown vertically upward will rise are found by setting the final velocity, $v = 0$, the value it assumes at the highest point. Then we have

$$t = \frac{v_0}{g} \tag{49}$$

$$s = \frac{v_0^2}{2g} \tag{50}$$

These values of t and s are the same as would be required to produce a velocity v_0, in the case of a body falling freely from rest.

24. Diminished Acceleration. Atwood's Machine. If we wish to study the laws of a falling body experimentally, it is necessary to reduce materially the acceleration, as otherwise the motion is much too rapid to permit of accurate observation. This may be done in several ways. For example, the force of gravity acting upon a small mass may be applied to one or more large masses as well, in which case the resultant acceleration is correspondingly diminished. Thus suppose that a body of mass M be placed upon a perfectly smooth, horizontal table and have attached to it a light flexible cord passing over a smooth peg at the end of the table, and that from this cord there be suspended a small mass m. In this case the stretching force in the string while the system is at rest will be that due to the weight of the mass m, and this force will produce motion in the two masses M and m. If we denote the resulting acceleration by a, we may equate the two expressions for the force as follows:

$$mg = (M + m)a \tag{51}$$

whence

$$a = \frac{m}{M + m} g \tag{52}$$

In this way the value of a may be made what we will, and the motion rendered so slow as to allow us to study it at leisure.

In *Atwood's machine* the light flexible cord passes over a light wheel having a groove in its rim and mounted upon "friction wheels" so as to turn as freely as possible. Two equal masses M_1 and M_2 are hung to the ends of the cord and are in equilibrium in any position. If now there be placed upon one of the masses a small rider of mass m, then the resultant motion of the system is due to the force of gravity upon this small mass alone. If we set M equal to the combined masses M_1 and M_2, and neglect friction and the effect of the light wheel, we may compute the resulting acceleration a at once from equation (52), and verify the result by actually observing the spaces passed over in one, two, three, or four seconds respectively. In case the rider be removed at any time, the acceleration becomes zero from that instant, and the motion becomes uniform motion. By means of special devices this may be accomplished and the machine may be used to verify all the conclusions represented in equations (48).

25. Motion on an Inclined Plane. The inclined plane is another device for reducing the effect of gravity. Suppose a particle of mass m to slide without friction down a plane AB (Fig. 8), making an angle ϕ with the horizon; it is required to find the equations of its motion. It is to be observed that the acceleration due to gravity effective in producing motion down the plane is the component parallel to the surface of the plane. This component is readily found by projecting g upon the plane (Art. 11). The effective component is seen to be $g \cos(90° - \phi)$ or $g \sin\phi$, and the equations of (47) become

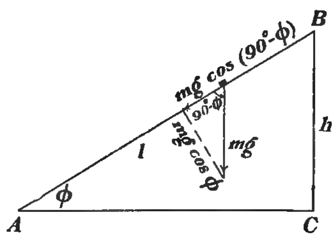


FIG. 8.

$$v = v_0 \pm g \sin\phi\, t$$

$$s = v_0 t \pm \frac{g \sin\phi\, t^2}{2} \tag{53}$$

$$v^2 = v_0^2 \pm 2\, g \sin\phi\, s$$

Substituting for $\sin\phi$ its value h/l, and setting $v_0 = 0$, and $s = l$, in the last equation of (53) we see that

$$v^2 = 2\,gh$$

or the velocity acquired by a body starting from rest and sliding down the plane is the same as that which it would have acquired falling through the vertical height h.

26. Uniform Circular Motion. A material particle describing circular motion at a constant speed is yet under the action of a constant force, since the *velocity* changes at every instant, not in *magnitude* but in *direction*.

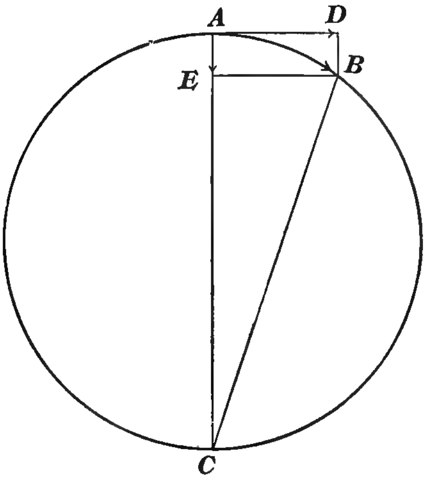


FIG. 9.

In uniform circular motion, therefore, there exists at every instant an acceleration toward the center. The value of this acceleration is readily calculated. Thus, in Fig. 9, a particle performing uniform circular motion describes the arc AB in time t, with a speed $\bar{v}$. It has in the same time been deflected from a straight line AD, through the distance AE, or AE is the space described due to the constant force acting toward the center. If $\bar{a}$ be the acceleration toward the center, then

$$AE = \frac{\bar{a}t^2}{2} \tag{54}$$

and

$$AB = \bar{v}t \tag{55}$$

since the motion is uniform.

If now the arc AB be taken very small, it will differ but little from a straight line, and the two triangles ABE and ABC are similar; hence

$$\frac{AB}{AE}=\frac{AC}{AB} \tag{56}$$

or

$$\overline{AB}^2=AE\cdot AC \tag{57}$$

whence

$$\bar{v}^2t^2=\frac{\bar{a}t^2}{2}\cdot 2r \tag{58}$$

or

$$\bar{a}=\frac{\bar{v}^2}{r} \tag{59}$$

If T be the time of a complete revolution of the particle in the circle, then

$$\bar{v}=\frac{2\pi r}{T} \tag{60}$$

and

$$\bar{a}=\frac{4\pi^2 r}{T^2} \tag{61}$$

whence

$$\bar{a}=\omega^2 r \tag{62}$$

27. Applications of Uniform Circular Motion. We have seen that in uniform circular motion a constant force is needed to keep the particle from flying off on a tangent. The force is termed *centripetal force*, and its measure is the product of the mass times the acceleration, or

$$F=m\frac{\bar{v}^2}{r} \tag{63}$$

The tendency of the particle to fly off, or the reaction of the particle against being pulled out of a straight line, is called *centrifugal force*, and its measure is also $m\bar{v}^2/r$. These two so-called forces are in reality but the two aspects of a *stress*, the one being the *action* of the moving mass, tending to move in a straight line, and the other the *reaction* of the center.

Examples of applications of these forces are seen in the action of a boy's sling; in the tendency of a bicycle rider to fall when turning a corner sharply while riding at full speed; and in the necessity for raising the outer rail in railway curves. The shape of the earth is an example of this action on the plastic mass of the earth while in rotation. Other applications

of this principle are seen in machines for drying clothes; for separating honey from the honeycomb, or molasses from sugar Cream separators, blood testers, lawn sprinklers, and the mechanical governors on steam engines all illustrate the same principle.

28. Simple Harmonic Motion. Fundamental Ideas. If from some solid support we suspend a spiral spring, in front of a mirror scale, and to the lower end of the spring attach a pan containing a small mass M (Fig. 10), we shall find that the system will soon come to rest at some point whose position may be accurately read off on the scale by means of the pointer. If now there be placed in the pan masses of 1, 2, and 4 grams each, and the successive readings of the pointer recorded, it will be found that the displacement of the pan from its position of rest for 4 grams is twice the displacement for 2 grams and four times the displacement for 1 gram. In short the displacement is directly proportional to the force applied. Again, since the restoring force of the spring exactly balances the weights applied to it, it follows that for a given elongation, the restoring force is proportional to the displacement of the moving system, and since the mass of this system is constant, it also follows that the acceleration due to this force must be proportional to the displacement.

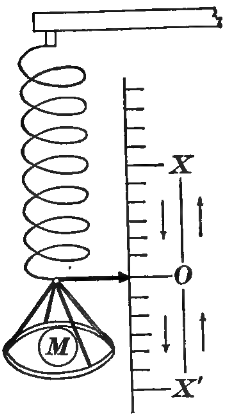


FIG. 10.

If, on a pair of rectangular axes XX' and YY' (Fig. 11), we let O represent the position of rest of the system, we may represent the downward displacements by spaces laid off on the negative axis of x, or OX', and upward displacements by corresponding spaces on OX. Then if x represent the displacement of the system from O *at any time*, and a_x be the acceleration along the axis of x, we may express the relations described in the previous paragraph by the equation

$$a_x = -c \cdot x \tag{64}$$

where the constant c represents the constant ratio between the acceleration a_x and the displacement x. The negative sign denotes that the *acceleration is always opposite in sense to the displacement.* Thus if the displacement of the spring be *downward* or *negative*, then the acceleration due to the recoil of the spring is *positive* or *upward*, and *vice versa.*

If now we pull down the system till the displacement x reach any value, say 10 cm, and then release it, we shall have a regular, periodic, vibratory motion in a vertical line, above and below the point O (Fig. 10). The system in its vibratory motion traces the path $X'OXOX'$ in the order indicated. Such a motion is a *simple harmonic motion. In simple harmonic motion the acceleration is proportional to the displacement from the position of rest and directed toward that position.*

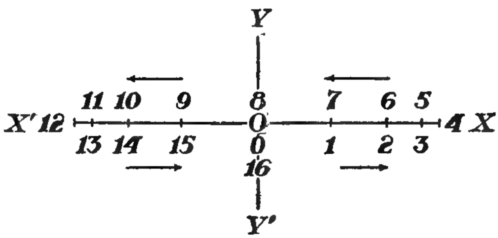


FIG. 11.

The maximum displacement of the system from the point O is called the *amplitude a*, of the vibration. The time elapsing between two successive passages through the same point in its path *in the same direction* is called the *period T* of the vibration.

If it were possible to take snap-shot photographs of the body, every $\frac{T}{16}$ sec during its entire vibration, the body would be seen at the points indicated by the figures on the axis, in the order shown by the arrows.

From the figure it is clear that the velocity of a point executing simple harmonic motion, is directed half the time in one direction and half the time in the other. Also that the velocity is a maximum at O, and zero at X and X'.

In order to locate completely at any instant of time a body describing simple harmonic motion it is necessary to know the time which has elapsed since the body passed some fixed point,

going in a definite direction. This elapsed time is usually counted from the instant the body passes the position of rest O, going in the positive direction, and is frequently measured in terms of the period.

Examples of simple harmonic motion are seen in the vibrations of the prongs of a tuning fork, of guitar strings or of a pendulum bob, if the amplitude be small; other examples of approximately this form of motion are found in the motion of the piston rod of an engine, of the shuttle of a sewing machine or of the sickle of a reaping machine. Simple harmonic motion is the most important type of motion to be studied, as it finds its applications in sound, light and electricity as well as in mechanics.

29. Circle of Reference and Definitions. Simple harmonic motion may also be regarded as the apparent motion of a point describing uniform motion in a circle, when viewed at a great distance from the circle and in the plane of the circle. The apparent motion of the moons of Jupiter is a simple harmonic motion. These little bodies revolve about the planet in orbits nearly circular, and their motion as seen from the earth is an oscillatory motion about the planet as a center.

For purposes of study, simple harmonic motion is usually treated in connection with the related case of uniform motion in a circle. Thus if we consider a point moving uniformly round a circle (Fig. 12), in the direction of the arrow, then the projections of the point upon any diameter of this circle will represent a simple harmonic motion, upon that diameter. The projection will describe a complete to-and-fro vibration upon the diameter, while the point on the circumference describes a complete revolution.

Thus while the point on the circle moves from o' through C', B', A', to X, its projection falls successively upon O, c, b, a, and comes to rest for an instant at X. It then retraces its path to O as the point passes round to o. The point on the circle has described half a circumference and its projection has made half a vibration. As the point passes on through D, E, F, to X', its projection swings through d, e, f, reaching the

limit of its motion in X', and returning completes the second half of its vibration as the point passes through F', E', D', back to o'. The circle upon a diameter of which the simple harmonic vibration is supposed to occur is called *the circle of reference.*

The *amplitude* of the vibration, that is, the maximum displacement from the center, now becomes the radius of the circle of reference. The *period* of the vibration, or the time required to make a complete to-and-fro vibration, now becomes the time required for the moving point to make a complete revolution in the circle.

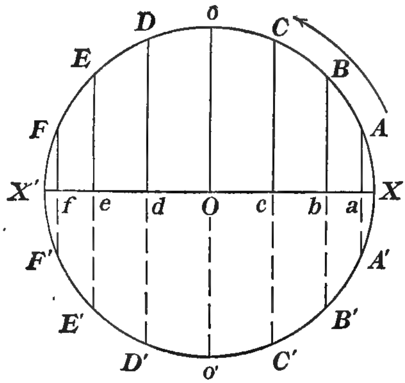


FIG. 12.

Phase is that fraction of a period which has elapsed since the moving point last passed through the position of rest in the positive direction. Phase may be expressed either in time, or in an angle which varies as the time.

30. Phase Relations. If ω represent the angular velocity of the radius vector CP (Fig. 13), then

$$\theta = \omega t$$

is the angle swept out in time t.

The angle θ is called the *time angle, i.e.* the angle swept out in time t, if time be counted from the instant the point P passes through X. In the case of motion on the axis of y, it is also the phase angle, since by definition, phase is measured from the instant at which the point P passes X, or its projection on the axis of y passes through C, in the positive direction.

For the same reason it is to be observed that in the case of motion on the axis of x (Fig. 13), phase must be measured from the radius CY', since the point P is at Y' when its projection s, on the axis of x, passes through C going in the posi-

tive direction. In this case the phase angle is $Y'CP$, or $\theta + 90°$. This shows that uniform circular motion is equivalent to two simple harmonic motions at right angles to each other, of the same period and amplitude, and differing in phase by 90°, or a quarter of a period.

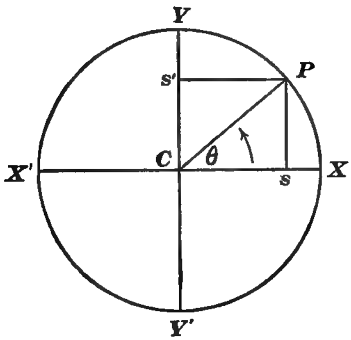


FIG. 13.

Again if we should begin to count time from some other point, as E (Fig. 14), then for the motion on the axis of y the time angle θ is ECP, while the phase angle is XCP as before, or $(\theta - \epsilon)$. For the motion on the axis of x the time angle is ECP, but the phase angle is $Y'CP$ as before, or $(\theta + \epsilon')$. The angles ϵ and ϵ' are called the *epoch angles*. The *epoch angle* is the angular difference between the time angle and the phase angle, or it is the angle which must be added to or subtracted from the time angle to produce the phase angle.

It is further to be noted that while ϵ and T are constant for any specific case, the time angle θ grows continuously from 0° to 360° while t grows from 0 to T.

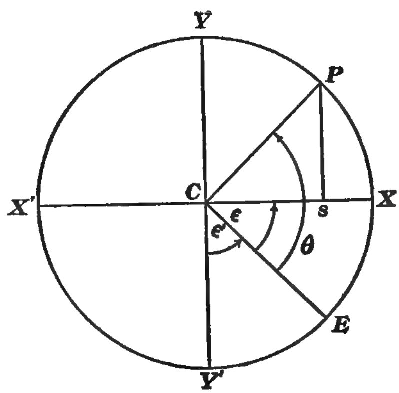


FIG. 14.

31. Equations of Simple Harmonic Motion. In order to describe in mathematical terms the behavior of a point executing simple harmonic motion it is customary to express its distance from the position of rest in terms of an angle that varies as the time. Thus if CX and CY (Fig. 15) represent the amplitudes of the two simple harmonic motions described by the two points s and s', the one

moving on the x-axis and the other on the y-axis, then these two separate motions are completely described if we write

$$\begin{aligned} x &= r \sin Y'CP \\ y &= r \sin XCP \end{aligned} \qquad (65)$$

or

$$\begin{aligned} x &= r \sin (\theta + \epsilon') \\ y &= r \sin (\theta - \epsilon) \end{aligned} \qquad (66)$$

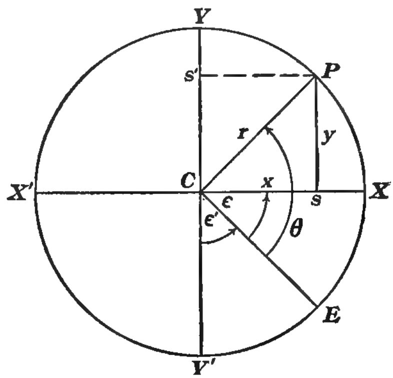


FIG. 15.

As a special case we may assume the point E to coincide with X, then $\theta = XCP$ $\epsilon = 0°$, and $\epsilon' = 90°$, and equations (66) become

$$\begin{aligned} x &= r \cos \theta \\ y &= r \sin \theta \end{aligned} \qquad (67)$$

Again, since the acceleration in uniform circular motion is

$$\bar{a} = \frac{4\,\pi^2 r}{T^2}$$

and *directed toward the center*, we have by projection upon the axes of x and y, the following expressions for the accelerations along these axes respectively:

$$\begin{aligned} a_x &= -\frac{4\,\pi^2 r}{T^2} \cdot \cos\theta = -\,\omega^2 r \cdot \cos\theta \\ a_y &= -\frac{4\,\pi^2 r}{T^2} \cdot \sin\theta = -\,\omega^2 r \cdot \sin\theta \end{aligned} \qquad (68)$$

The negative sign denotes that the acceleration is always opposite in sign to the displacement. Equations (68) show that the acceleration along the axis of x is a maximum when that along the axis of y is 0, and *vice versa*, since the one varies as the sine and the other as the cosine of the same angle.

If we substitute in equations (68) the values of $\sin\theta$ and $\cos\theta$, we have

$$\begin{aligned} a_x &= -\frac{4\,\pi^2 r}{T^2} \cdot \frac{x}{r} = -\frac{4\,\pi^2 \cdot x}{T^2} = -\,\omega^2 x \\ a_y &= -\frac{4\,\pi^2 r}{T^2} \cdot \frac{y}{r} = -\frac{4\,\pi^2 \cdot y}{T^2} = -\,\omega^2 y \end{aligned} \qquad (69)$$

which shows that the acceleration along the the axis of x is proportional to x, and that along the axis of y is proportional to y, or that *the acceleration is proportional to the displacement.* This is the characteristic of simple harmonic motion.

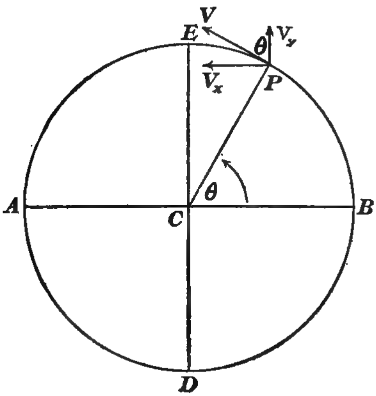


FIG. 16.

32. Velocity of a Point executing Simple Harmonic Motion. The velocity of a point executing simple harmonic motion may be readily calculated. Let P (Fig. 16) be the moving point, and let PV represent its velocity in the circle. Then V_x and V_y, the component velocities parallel to the axes of x and y, are found by projection as usual, and we have

$$V_x = -V\cos(90° - \theta) = -V\sin\theta \qquad (70)$$
$$V_y = V\cos\theta$$

But V is the velocity in the circle, and is therefore $\frac{2\pi r}{T}$, hence

$$V_x = -\frac{2\pi r}{T}\cdot\sin\theta \qquad V_y = \frac{2\pi r}{T}\cdot\cos\theta \qquad (71)$$

From equations (71) it appears that the velocity along the axis of x is a maximum when the velocity along the axis of y is zero, and *vice versa.* Also from comparison with (68) we see that when the velocity along the axis of x is a maximum, the acceleration along x is zero; this means the velocity is greatest at the middle of the swing where the acceleration is zero, and zero at the end of the swing where the acceleration is greatest.

It should also be noted that since the angular velocity ω is a constant, the time of a vibration, T, is also a constant, and the vibrations of a body executing simple harmonic motion *are all performed in the same time, and are independent of the ampli-*

tude. Such vibrations are said to be *isochronous*. This is seen in the constancy of the pitch of the musical note from a string or tuning fork, as the vibrations die away. This characteristic of simple harmonic motion is of high importance in the theory of the pendulum, in acoustics and optics.

33. The Curve of Sines. An important aid to the study of simple harmonic motion is found in the graphical method, whereby the moving body is made to trace its path upon some recording surface. In such cases the simple harmonic motion

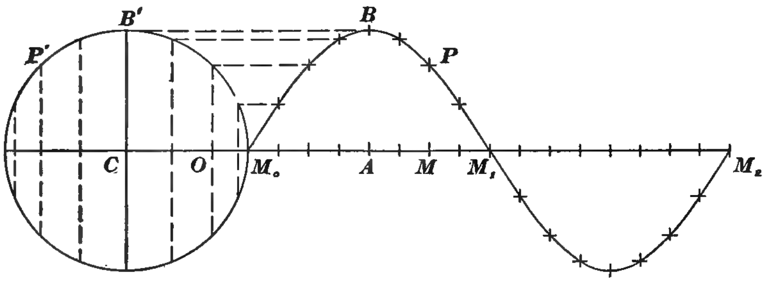


FIG. 17.

is compounded with a uniform motion in a straight line and the resultant curve is called a *sine curve*. Such a curve is readily constructed as follows: Let the simple harmonic motion be described on the vertical axis CB' about C as a center (Fig. 17). Divide each quadrant of the circle of reference into four equal parts, each of which will thus correspond to the distance passed over by the point moving in the circle, in one sixteenth of a period. From the points of division drop perpendiculars to the x diameter; these perpendiculars will then denote the displacements of the point executing simple harmonic motion on the y axis, at the ends of the successive intervals of time, and are therefore proportional to the sines of the angles swept out in $T/16$, $2T/16$, $3T/16$, etc., beginning with the instant when the moving point passes through C, going in the positive direction. Next lay off on the x diameter produced a series of *equal lengths* to represent the spaces described in the *same intervals*,

due to the *uniform motion*, and beginning at M_0, erect perpendiculars equal to the y displacements at the corresponding times. Finally sketch a smooth curve through the extremities of these perpendiculars and the *sine curve* is the result.

Such a curve is readily obtained by allowing a tuning fork moving uniformly to trace, with a fine style, its vibrations upon a sheet of smoked glass or paper. The time interval between M_0 and M_1 corresponds to a half period, or the points at M_0 and M_1 *differ in phase* by half a period, while the points at M_0 and M_2 are in the the *same phase*.

The distance M_0M_1 is a *half wave length*, and M_0M_2 is a *complete wave length*, where a *wave length* denotes the distance the wave form has run forward in a single period T. Hence we have the important relation

$$\lambda = VT \tag{72}$$

where λ denotes the wave length, V the velocity of the wave motion, and T the period of the simple harmonic vibration.

Problems

1. A body has a velocity of 60 mi an hour. Find its velocity in feet per second, and in centimeters per second. *Ans.* 88 ft per sec. 2684 cm per sec.

2. A body starts with a velocity of 640 cm per second, and in 10 min has a velocity of 3428 cm per second. Find the acceleration.

Ans. $a = 4.61 \dfrac{\text{cm}}{\text{sec}^2}$.

3. A mass of 5000 g moves with a velocity of 4 m per second. Find its momentum.

Ans. $2 \times 10^6 \text{ g} \dfrac{\text{cm}}{\text{sec}}$.

4. A mass of 600 g starts from rest and in 6 sec has a velocity of 36 m per second. Find the force. *Ans.* $F = 36 \times 10^4$ dynes.

5. A vessel contains 400 cc of sulphuric acid, density 1.8 g per cubic centimeter. Find the mass of the acid. *Ans.* 720 g.

6. Detroit is 38 mi from Ann Arbor. How long will it take to travel this distance at the rate of 5 km per hour? *Ans.* 12.23 hr.

7. Express a mass of 65 lb in grams; a weight of 65 lb in dynes.

Ans. 29545 grams. 28.954×10^6 dynes.

8. What force will in 10 min give a mass of 700,000 g a velocity of 124,000 cm per second? *Ans.* $F = 1446.7 \times 10^5$ dynes.

9. What is the force of gravity on a body whose mass is 700 lb? *Ans.* 311.8×10^6 dynes.

10. Find the mass of 465 cc of lead. Density of lead = 11.3 g per cm^3. *Ans.* 5254.5 grams.

11. A wire guy rope, the stretching force on which is 12×10^8 dynes, makes an angle of 60° with the horizon. Find the vertical and the horizontal components of the stretching force. *Ans.* $F_h = 6 \times 10^8$ dynes.
$F_v = 10.392 \times 10^8$ dynes.

12. Find the acceleration produced upon a mass of 4 g by a force of 36 dynes. *Ans.* $a = 9 \dfrac{\text{cm}}{\text{sec}^2}$.

13. A force of 60 dynes acts upon a body for one minute and imparts to it a velocity of 900 cm per second. Determine the mass of the body. *Ans.* 4 g.

14. An engine draws a cage of mass 3000 kilos up a shaft at a uniform speed of 10 m per second. (*a*) Find the stretching force in the rope. (*b*) What is the stretching force if the cage move with a uniform acceleration of 10 m per second per second? *Ans.* (*a*) 294×10^7 dynes.
(*b*) 594×10^7 dynes.

15. An elevator starts to descend with an acceleration of 300 cm per second per second. (*a*) Find the apparent weight on its floor due to a man whose mass is 75 kilos. (*b*) What would be his weight with respect to the elevator, if it started to *ascend* with the same acceleration?
Ans. (*a*) 51×10^6 dynes.
(*b*) 96×10^6 dynes.

16. A steamer whose velocity in still water is 6 mi an hour starts directly across a stream whose velocity is 10 mi an hour. Find the velocity of the steamer in crossing; also when it makes an angle of 30° with the current down stream. *Ans.* (*a*) 11.65 mi per hour.
(*b*) 15.48 mi per hour.

17. A gun of mass 3000 kilos, placed upon a smooth horizontal plane, discharges a ball of 30 kilos mass at an elevation of 30° to the horizon. Find the velocity of the gun's recoil in terms of velocity of the ball.
Ans. $0.00866\ V_b$.

18. An inelastic mass of 900 kilos moving with a velocity of 30 m per second meets another equal and similar mass moving in the opposite direction, at 10 m per second. Find velocity of total mass after impact.
Ans. 10 m per sec.

19. A certain force acts upon m units of mass, and at the end of a second the mass is moving at the rate of 32 ft per second. What velocity would be produced in 32 units of mass by the same force, in the same time?

Ans. m ft per sec.

20. How many dynes are required to give a mass of 50 kilos a velocity of 12 m per second, the force being supposed to act for exactly one second?

Ans. 6×10^7 dynes.

21. How many dynes are needed to give a gram a velocity of 9.81 m per second, if the force act for one second? What if it act for two seconds?

Ans. (*a*) 981 dynes.
(*b*) 490.5 dynes.

22. A body falls freely from rest for 15.6 sec. Find the final velocity and the distance traversed. *Ans.* $v = 15{,}288$ cm per second.
$s = 119{,}246$ cm.

23. How long will it take a body to fall 650 ft, and what velocity will it acquire? *Ans.* (*a*) 6.35 sec.
(*b*) 6231 cm per sec.

24. A body is thrown downward with a velocity of 874 cm per second. Required its velocity and position at the end of 20 sec.

Ans. (*a*) 20,474 cm per sec.
(*b*) 213,480 cm below starting point.

25. A body is thrown vertically upward with a velocity of 827 cm per second. How long will it continue to rise, and how high will it rise?

Ans. (*a*) 0.84 sec.
(*b*) 348.94 cm.

26. A body is thrown vertically upward with a velocity of 697 cm per second. When will it be 195 cm above the starting point, and what velocity will it then possess? *Ans.* (*a*) 0.383 sec or 1.039 sec. Explain the two values for t. (*b*) 321.66 cm. per sec.

27. A ball is thrown vertically upward to a height of 150 ft. With what velocity did it leave the hand? ($g = 32.16$ ft/sec^2.) *Ans.* 98.22 ft per sec.

28. A mass of 876 g is attached to a spring balance which is carried upward at such a rate that the balance indicates 932 g. What is the acceleration of the motion? *Ans.* 62.6 cm/sec^2.

29. A mass of 162 kilos hanging by a perfectly flexible cord drags a mass of 973 kilos along the top of a smooth table. What is the acceleration of the system, and what is the stretching force in the cord?

Ans. $139.89 \dfrac{\text{cm}}{\text{sec}^2}$.

Stretching force in cord $= 136.1 \times 10^6$ dynes.

30. Two masses of 100 g each are hung by a flexible cord over a frictionless pulley. A mass of 10 g is placed upon one of the 100 g masses Required the acceleration of the system and the stretching force in the cord.

Ans. $46.66 \frac{\text{cm}}{\text{sec}^2}$. Stretching force = 102,666 dynes.

31. Masses of 938 and 762 g respectively are hung by a flexible cord over a frictionless pulley. How far must the masses move in order to acquire a velocity of 325 cm per second? *Ans.* 520.53 cm.

32. A body slides down a smooth plane 326 cm long, inclined at an angle of 45° to the horizon. Find the time of descent and the velocity with which it reaches the bottom. *Ans.* 0.970 sec.
679.6 cm per sec.

33. A mass of 200 g is constrained to move in a circle of 600 cm radius with a velocity of 240 cm per second. What is the centripetal force and the period of revolution? *Ans.* (*a*) 19,200 dynes.
(*b*) 15.708 sec.

34. The distance of the moon from the earth is 3.84×10^{10} cm, and the lunar month is approximately 27 da and 8 hr. What is the acceleration due to the earth's attraction at the moon?

Ans. $a = 0.2727 \frac{\text{cm}}{\text{sec}^2}$.

35. If a skater describe a circle of 100 ft radius with a speed of 20 ft per second, find the inclination of his body from the vertical in order that he may maintain his equilibrium. *Ans.* 7° 5′ 30″.

36. If the equatorial radius of the earth is 3963.3 mi, find the time of rotation necessary for a body at the equator to weigh nothing, assuming $g = 981$ cm per second per second for the earth at rest. *Ans.* 1 hr 26 min.

37. A mass of 1 g moves uniformly round a circle 40 cm in diameter at the rate of 24 revolutions a minute. Compute the force toward the center. *Ans.* 126.23 dynes.

38. A mass of 1 g executes simple harmonic motion with an amplitude of 4 cm and a period of 0.5 sec. Find the force toward the center when the phase is $T/4$, $T/8$, and $T/2$, respectively.

Ans. $64\pi^2$ dynes; $45.25\pi^2$ dynes; 0.

39. Show that the spaces passed over, starting from the center, by a body executing simple harmonic motion, in the successive time intervals $T/16$, are approximately proportional to 4, 3, 2, and 1.

40. A horizontal shelf moves vertically with simple harmonic motion, with a period equal to one second. Find the maximum amplitude it can have so that objects resting upon it may remain in contact with it at its highest point. $g = 980$ cm per second per second. *Ans.* 24.82 cm.

CHAPTER IV

WORK AND ENERGY

34. Work. Work consists in changing a state of motion or a state of stress, in opposition to forces tending to resist such an effect.

Examples of work are seen (*a*) in the starting of a heavy car upon a smooth, level track; (*b*) in the compressing of a gas into a cylinder; (*c*) in the pumping of water into an elevated reservoir; (*d*) in the winding of a watch; (*e*) in the charging of a storage battery; (*f*) in the action of gravity upon a system composed of a wheel and axle, to which is attached a heavy weight by means of a cord wound round the axle.

When a body moves in the direction of the force acting upon it, the force is said to do work upon the body in giving it motion, as in the case of a freely falling body; if the motion be in opposition to the force, work is said to be done upon the body *against* the force, in giving the body a change in condition, or putting it in a state of stress; an example of this is seen in compressing gas into a cylinder.

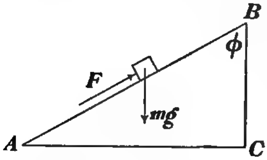


FIG. 18.

The measure of work done is the product of the *force* into the *distance*, in the direction of the force, or

$$W = Fs \tag{73}$$

If the motion take place in a line inclined to the direction of the force worked against, the effective displacement is found by projection, and the work is the product of the force into the effective displacement. An example of this is the work done

against gravity in moving a body up an inclined plane. Thus the force F (Fig. 18), in working over the distance $s = AB$, is, in reality, lifting the mass through the distance CB. The effective displacement is therefore the vertical component of s, or $s \cos \phi$, and in this case the work is

$$W = mgs \cos \phi \tag{74}$$

therefore

$$F = mg \cos \phi \tag{75}$$

where ϕ is the angle between the direction of the motion and that of the force.

The C. G. S. unit of work is the *erg*. An erg is the work done by a force of one dyne acting through a distance of one centimeter; or unit work is done by unit force acting through unit distance. The force exerted by gravity upon a gram mass is 980 dynes. Therefore, to lift a gram mass one centimeter against gravity would require 980 ergs.

Since the erg is a very small unit, the joule $= 10^7$ ergs is generally used.

The dimensions of work are

$$[MLT^{-2} \times L] = [ML^2T^{-2}].$$

In the gravitational system of units the unit of work is either the foot-pound or the kilogram-meter. The foot-pound denotes the work necessary to lift a pound mass through a distance of one foot against the force of gravity, while the kilogram-meter denotes the work done in raising a mass of one kilogram through a distance of one meter against gravity.

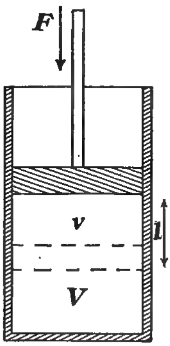


FIG. 19.

35. Work done by a Gas expanding under Constant Pressure. Let a mass of gas (Fig. 19) be inclosed in a cylinder, furnished with a frictionless piston of area a and mass m. Equilibrium will be established when the internal pressure of the gas just balances the external pressure, or when the upward force due to the stress in the gas equals the weight of the atmosphere upon the piston plus the weight of the piston. This force is

$$F = Ba + mg = Pa \tag{76}$$

where B is the atmospheric pressure, and P the pressure on the gas.

If now the gas be heated, it expands and pushes the piston forward through a distance l. The work done by the gas is therefore

$$Work = Fl = Pal = Pv \tag{77}$$

where v is the change of volume produced.

The work done *by* a gas, expanding under *constant* pressure, is thus equal to the product of the pressure and the increase in volume. Conversely, if a constant pressure P produce a decrease v in the volume of a gas, the product Pv is the work done *upon* the gas.

In general the pressure of a gas is not constant while it is being compressed, but increases with the compression. In this case the above equation can be applied only for very small compressions during which the pressure may be assumed to remain constant.

36. Power. It is to be noted that the expression for work

$$W = Fs$$

contains no element of time. The same work is done in lifting a bag of grain to the top of a building, whether the work be done in an hour or in a month. *Power* involves the idea of the *rate at which work is done*, and may be defined as the *time rate of doing work.* Hence

$$Power = \mathcal{P} = \frac{W}{T} = \frac{Fs}{T} = \frac{Pv}{T} \tag{78}$$

The dimensions of power are $[ML^2T^{-3}]$.

The C. G. S. unit of power is the *watt.* A watt is the power which will do 10,000,000 ergs, or 1 *joule*, in one second. The practical unit of power is the *kilowatt*, or one thousand watts. The English unit of power is the *horse power*, which denotes the power to do 33,000 foot pounds of work per minute, or 550 foot pounds per second.

One horse power is equal to 746 watts.

37. Energy. Whenever work has been done upon a system in producing a *change* either in its *motion*, its *position* or its *molecular condition*, the system has acquired the capability of doing work in turn.

Energy is the capability of doing work, possessed by a system by virtue of work having been previously done upon it. This is seen to be true of all the cases already cited (Art. 34). The car, once set in motion, can do work by virtue of that motion and to the extent of that motion; the wheel and axle and the falling weight each possesses energy, the one by virtue of its motion of translation, and the other by virtue of its motion of rotation; the gas compressed in the cylinder or the coiled spring of a watch each possesses energy due to a *stress* or a tendency to return to a former state. The water in the reservoir possesses energy due to gravitational stress; the chemical elements in a storage battery, having been separated by the electric current, now tend to reunite and thus possess energy due to chemical stress.

Energy is thus seen to exist in one of two distinct forms:

(*a*) **Energy of Motion,** *or Kinetic Energy;*
(*b*) **Energy of Stress,** *or Potential Energy.*

If we do work upon a system, we increase its energy, since we transfer energy from our bodies to the system. The sum total of energy in the system working and the system worked upon is at all times a constant quantity. When work is done by one system upon another, both kinds of energy are present, and there is a transfer of energy from the system working to the system worked upon. Hence to do work is to transfer energy from one system to another.

It is to be noted that whenever the motion is *with* the force, the motion of the system is increased and the system gains in kinetic energy, as in the case of a freely falling body. When the motion is *against* the force, kinetic energy is changed into potential, as in the case of a body thrown vertically upward.

Strictly speaking the system working is able to transfer only its available energy to the system worked upon. The working

system may possess energy in the form of heat that is unavailable for the purpose of doing useful work. In the same way the system worked upon may receive energy from the system working in the form of heat, which escapes later, and is no longer available as energy in the system worked upon.

38. Expressions for Energy. Since by definition a system possesses energy only by virtue of work done upon it, the unit of energy is the same as the unit of work, *the erg*. The dimensions of energy are the same as those of work, $[ML^2T^{-2}]$.

Potential energy is stored up work, and its expression is

$$W = Fs = Mas = \textit{Potential Energy} \tag{79}$$

To express the kinetic energy of a body in terms of its mass and velocity, we need to remember that we have a force F, acting through a space s, and the expression for kinetic energy may be deduced in a number of ways. Thus, if a mass m be acted upon by a force F, for a time t, during which it receives an acceleration a, it will pass over a space

$$s = \tfrac{1}{2}\, at^2$$

and acquire a velocity

$$v = at$$

Then the work done by the force will be

$$W = Fs = F \cdot \frac{at^2}{2} = \frac{ma^2t^2}{2} = \frac{mv^2}{2} \tag{80}$$

or

$$\textit{Kinetic Energy} = \frac{mv^2}{2} \tag{81}$$

Again, since

$$v^2 = 2\, as$$

we may write

$$mv^2 = 2\, mas \tag{82}$$

or

$$\frac{mv^2}{2} = mas = Fs = \textit{Kinetic Energy}$$

Finally we may deduce the same expression from fundamental definitions. Thus, the impulse is equal to the force into the time during which the force acts, or

$$Impulse = Ft = mv - mv_0 \tag{83}$$

also the mean velocity v', is

$$v' = \frac{s}{t} = \frac{v + v_0}{2} \tag{84}$$

whence, combining (83) and (84),

$$Fs = \frac{mv^2}{2} - \frac{mv_0^2}{2} \tag{85}$$

This last expression states that *the work* is equal to the *change in kinetic energy* produced, and is a more general expression for the work done upon a body. By setting v_0 equal to 0, we have

$$K.E. = \frac{mv^2}{2}$$

as before. Since kinetic energy like work is a scalar quantity and therefore independent of direction, we may substitute speed for velocity in the above formula.

39. Transformations of Energy. Transformations of energy occur on every hand. An excellent example is seen in the motion of a pendulum bob. At the highest point of the swing its energy is all potential, at the lowest point it is all kinetic; at intermediate points it is partly kinetic and partly potential. Were it not for the slight loss of energy in overcoming the resistance of the air and the stiffness of the cord, this transformation would go on forever.

Consider also the transformations of energy presented in the consumption of coal in the furnace of a steam boiler. The coal supply of the world represents the largest available source of potential energy and is simply the stored up sunshine of geologic ages. When the coal is burned in the furnace, this energy becomes kinetic in the form of heat, it appears as kinetic energy in the molecular motions of the water particles in steam, and as potential energy whose effect is an increase of steam pressure on the boiler. In the steam engine the energy becomes the kinetic energy of the moving masses of the machinery and

belts, which in turn may be transformed into light, motion, electric current or energy of chemical stress, and finally into heat again.

The transformation of energy from the potential to the kinetic form is *always a perfect one*, in that all the potential energy appears as kinetic. On the other hand the transformation from kinetic to potential is *never perfect;* some of the energy escapes as diffused heat and thus becomes unavailable for the purpose of doing useful work.

Potential energy tends to become a minimum. If in any system, any one of the stresses acting be removed, a redistribution of the energy occurs and the potential energy diminishes while the kinetic energy increases. An example of this is seen in the bursting of a reservoir full of water. The reaction of the restraining wall being removed, the water is carried down hill with increasing velocity by the force of gravity. The kinetic energy is increased at the expense of the potential. Other illustrations are seen in the bursting of a soap bubble, in the concentration of a dewdrop into a sphere, and in the position assumed by any body free to move into a new position of equilibrium. The result in all cases is that the system is at rest only when the potential energy is as small as possible.

40. Conservation of Energy. Throughout all the various transformations of energy it is to be noted that no body or system of bodies can acquire energy save at the expense of energy possessed by some other system. Hence we may say that to do work is to transfer energy from one system to another, and it seems certain, from the most careful experiments, that *the amount of energy lost by the one system is the exact equivalent of that gained by the other.* This means that no machine or combination of machines can ever be made to return more energy than is given to it. Perpetual motion is a delusion. Physically it is impossible to get something for nothing. Everything must be paid for in terms of energy.

Not only is it impossible to make a machine that will create energy, but no machine will ever return all the energy put into it. Owing to friction between the parts of the machine, some

energy is transformed into heat and rendered unavailable for doing useful work. It is not lost, since energy, like matter, is indestructible.

The doctrine of the conservation of energy states that *in a system so situated that it neither loses energy from within nor gains energy from without, the amount of energy is constant.* No energy can be created, none can be destroyed.

Problems

1. How much work in ergs will be done by a force of 48 dynes acting through 24 cm? *Ans.* 1152 ergs.

2. What work will be required to lift 10 kg of water from a well 12.5 m deep? *Ans.* 1225×10^7 ergs.

3. The lower end of a ladder 16 m long stands on the ground at a distance of 273 cm from a wall against which the upper end rests. How much work will be done in carrying 30 kilos up the ladder? *Ans.* 4634.91 joules.

4. The diameter of the cylinder of a steam engine is 18 in and its length is 24 in. What work in foot pounds will be done at each stroke of the piston if the average pressure of the steam be 110 lb per square inch? *Ans.* 55,983.3 foot-pounds.

5. If the above engine make 100 strokes per minute, calculate the horse power it will develop. *Ans.* 169.64 H. P.

6. A shot of mass 2 kilos moving with a velocity of 200 m per second is just able to pierce a plank 4 cm thick. What velocity is required to pierce a plank 12 cm thick? *Ans.* 346.4 m per sec.

7. A stone of mass 5 kilos is thrown vertically upward with a velocity of 25 m per second. Find its kinetic energy at the end of two seconds. *Ans.* 729×10^6 ergs.

8. A bullet of 100 g mass is discharged from a gun of mass 3 kilos, with a velocity of 400 m per second. Compare the kinetic energies of bullet and of gun. *Ans.* As 30 to 1.

9. A ball of 25 kilos mass moves with a velocity 4 m per second. Compute its kinetic energy. *Ans.* 2×10^9 ergs.

10. A man whose mass is 160 lb carries a hod and mortar of mass 75 lb from the ground to a scaffold 24 ft high, every 10 min. At what rate is this work done? *Ans.* 564 ft lbs per min.

11. A standpipe 20 m high and 4 m in diameter is to be filled with water from a lake 8 m below the base of the standpipe. How long will it take a 10 H. P. engine to fill it? *Ans.* 1 hr 39 min.

12. A 100 gram mass is suspended from a spring balance which is carried in a balloon. What will be its apparent mass as shown by the index (*a*) when the balloon is ascending with a uniform acceleration of 240 cm per second per second? (*b*) when it is descending with an acceleration of 900 cm per second per second? *Ans.* (*a*) 124.49 g.
(*b*) 8.16 g.

13. Two unequal masses are attached to the ends of a cord passing over a smooth peg. Find the ratio between them in order that they may move through 500 cm in two seconds, starting from rest. *Ans.* As 1230 to 730.

14. The upper end of a smooth straight wire of length 100 ft, is attached to a pole 48 ft high. A bead is allowed to slip along the wire from top to bottom. Find its velocity on reaching the bottom. Also time of descent.
Ans. (*a*) 55.56 ft per sec.
(*b*) 3.59 sec.

15. A ball thrown up is caught by the thrower 7 sec afterwards. How high did it go, and with what speed was it thrown? How far below its highest point was it 4 sec after the start? *Ans.* (*a*) 6002.5 cm.
(*b*) 3430 cm per sec.
(*c*) 122.5 cm.

16. The mass of a pendulum bob is 100 g, and the string is 1 m long. The bob is held so that the string is horizontal, and then allowed to fall. Find its kinetic energy when the string makes an angle of 30° with the vertical. *Ans.* 8.5×10^6 ergs.

CHAPTER V

MECHANICS OF A RIGID BODY

41. Motion of a Rigid Body. A rigid body is one that suffers no change of form as a result of the forces acting upon it. When force is applied to a rigid body that is free to move, the body will acquire motion of translation or of rotation, or of both together. The motion of translation imparted to the body is fully accounted for by the equation

$$F = Ma$$

which shows that the acceleration imparted to the body will vary directly as the force, and inversely as the mass of the body.

If the angular velocity of a body change, this change must be attributed to the action of a force. Angular acceleration is the *time rate of change of angular velocity.* If this angular acceleration α be constant, then, by definition, we have

$$\frac{\omega - \omega_0}{t} = \alpha = constant \tag{86}$$

If now a definite force be applied to a rotating body at any point, and in any direction, provided the force do not pass through the axis of rotation, the resulting angular acceleration will vary greatly, according to the direction of the force and the distance of its point of application from the axis of rotation. Thus, in order to produce rotation, a force must have a component normal to the direction of the axis, and the farther its point of application is from this axis, the greater will be the angular acceleration produced. This change in angular velocity is due to the action of what is called the *torque* $\mathcal{T}$, or the moment of the force about the axis. Hence the torque, or moment of a force about any given axis, is that which changes, or tends to change, the state of rotation of

the body with respect to this axis. This torque is the product of the component of the force normal to the axis into the perpendicular distance from the axis of rotation to the line of action of the force. Thus (Fig. 20) the moment of the force EF about the axis through O, normal to the plane of the paper, is $EF \times OP$. Torques are positive or negative, according as they tend to produce rotation in the counter-clockwise, or clockwise sense.

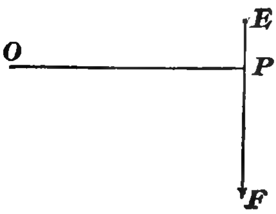


FIG. 20.

42. Resultant of Two Parallel Forces. Let F_1 and F_2 (Fig. 21) be two parallel forces applied to a rigid body in the directions $a'a$ and $b'b$; it is required to find the point of application C of the resultant R, such that its torque shall be equal to the sum of the torques of F_1 and F_2, or that it shall have the same effect in producing rotation about any point O as the combined effect of F_1 and F_2 taken together. From O drop a perpendicular upon the direction of F_1, F_2, and R, cutting them in A, B, and C. Represent OA, OB, and OC by x_1, x_2, and X respectively. Then, by the conditions of the problem,

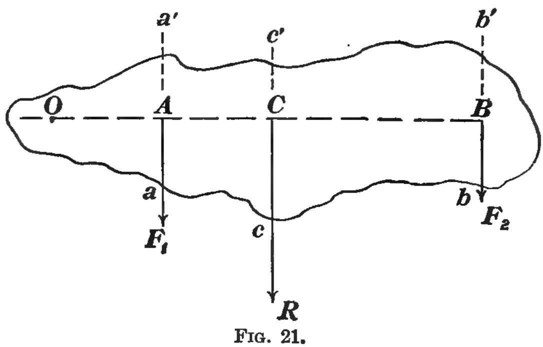


FIG. 21.

$$RX = F_1x_1 + F_2x_2 \tag{87}$$

But, by definition,

$$R = F_1 + F_2 \tag{88}$$

whence

$$X = \frac{F_1x_1 + F_2x_2}{F_1 + F_2} = \frac{\Sigma Fx}{\Sigma F} \tag{89}$$

or

$$\frac{F_1}{F_2} = \frac{x_2 - X}{X - x_1} \tag{90}$$

This shows that the point of application of the resultant divides the line joining the points of application of the two forces into two parts inversely as the forces. If the point O coincide with C, then

$$\frac{F_1}{F_2} = \frac{x_2}{-x_1} \tag{91}$$

or

$$F_2 x_2 = -F_1 x_1 \tag{92}$$

which means that the moments of F_1 and F_2 about C are equal and opposite, and there is, therefore, no tendency to rotate about this point, or the torques about C are equal and opposite.

If F_1 be equal and opposite to F_2, then the value for X, from (89), becomes infinite, or the solution fails for the case of *forces equal, parallel,* and *opposite in sense.* This simply means that *there is no point* at which a single force can be applied so as to produce equilibrium for this system. Such a system of forces is called a *couple.* *Two equal, parallel, and opposite forces constitute a couple,* and if applied to the adjacent parts of a rigid body, tend to produce rotation only. The moment of a couple is obviously the product of one of the forces into the perpendicular distance between them.

The only way by which a couple can be equilibrated is by means of another couple, of equal moment and oppositely directed.

43. Center of Inertia. If we take a body, as a long thin rod (Fig. 22), then every particle of matter in the rod is subject to the force of gravity, and the directions of these forces may be considered as sensibly parallel. If we take moments about any point O, situated somewhere on the axis of the rod, then by equation (89), we find for X, the acting distance of the resultant R,

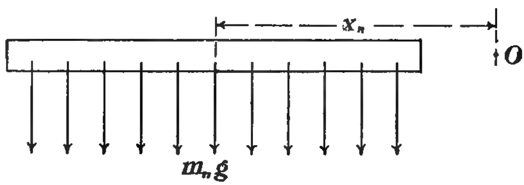


FIG. 22.

$$X = \frac{m_1gx_1 + m_2gx_2 + m_3gx_3 + \cdots + m_ngx_n}{m_1g + m_2g + m_3g + \cdots + m_ng} \tag{93}$$

or

$$X = \frac{\Sigma mgx}{\Sigma mg} = \frac{\Sigma mx}{\Sigma m} \tag{94}$$

In general, if x_1, x_2, $\cdots$ denote the perpendicular distances of the mass particles from any plane not parallel to the rod, then the length X determines the perpendicular distance from this plane to another plane intersecting the rod. The point of intersection is called the *center of inertia* or the *center of mass*, of the rod. It denotes the point of application of the resultant of all parallel forces of gravity acting upon the rod. If a rigid support be placed under this point in the rod, all the forces of gravity acting upon the rod will be in equilibrium, since the sum of their moments about this point is equal to zero; therefore this point is frequently called the *center of gravity* of the rod.

Equation (94) also affords a means for determining the center of figure of any plane area, A, made up of a large number of small areas, a_1, a_2, a_3, to a_n. If instead of mgx we substitute agh, and for mg we write ag, then H, the perpendicular distance from any plane of reference, not parallel to the area, to a plane intersecting the given area in a line passing through the center of figure is given by the equation

$$H = \frac{\Sigma agh}{\Sigma ag} = \frac{\Sigma ah}{\Sigma a} = \frac{\Sigma ah}{A} \tag{95}$$

By choosing a second plane of reference not parallel to the first, we may find as before, the perpendicular distance H', to a plane intersecting the area in a second line, also passing through its center of figure. The center of area must lie at the intersection of these two lines.

Similarly the center of mass of any solid may be found by using three planes of reference not parallel to one another. These planes of reference determine by equation (94) three other planes which intersect at the center of mass of the solid.

This general principle may be applied to a few simple cases.

(*a*) The center of inertia of a line is at its middle point.

(*b*) The center of inertia of any lamina having an axis of symmetry lies on this axis. If the lamina have two axes of symmetry, then the center of inertia of the lamina lies at their intersection.

(*c*) The center of inertia of a triangle lies at the intersection of the median lines of the triangle.

(*d*) The center of inertia of any polygon may be determined by dividing the polygon into triangles and considering the weight of the triangles as applied at their respective centers of inertia.

(*e*) The center of inertia of a triangular pyramid may be shown to be three fourths of the distance from the apex to the center of inertia of the triangular base of the pyramid.

44. Conditions of Equilibrium. Since the motion of a rigid body is either a motion of translation or of rotation, or a combination of the two, it follows that for the body to be in equilibrium, there should be no tendency to produce either translation or rotation. Hence the conditions for equilibrium are

$$\Sigma Fx = 0 \text{ (no rotation)} \tag{96}$$

$$\Sigma F = 0 \text{ (no translation)} \tag{97}$$

where moments and forces are to be taken with their proper signs. The origin or center about which the moments are to be taken may be situated anywhere, as its position is of no importance.

***45. Stability of Bodies.** The principle of moments may be applied to the determination of the stability of bodies. A body is said to be *stable*, or *in stable equilibrium*, when it is necessary to do work upon it in order to overturn it. In such a case the center of inertia must be raised, if the body is to be displaced, and the body, if disturbed, tends to return to its original position. If, on the other hand, the body on being disturbed has its center of inertia lowered, or tends to fall farther from its original position, it is said to be in *unstable*

equilibrium. A book lying on its side on a smooth table is in stable equilibrium. If the book be set upon edge or upon end, it is in relatively less stable equilibrium. A ball or a cone lying on its side upon a smooth horizontal surface is said to be in *neutral* or *indifferent equilibrium.*

The stability of quadrupeds is great, on account of the broad base, represented by the area enclosed by the four feet, and because the center of inertia is almost vertically over the center of the base. For the kangaroo, though an apparent biped, the tail acts as a third leg, and the base is correspondingly increased. In man the base of support is small and the center of inertia is high, and the stability of the upright body is not great. In corpulent persons the increased effort to keep the center of gravity over the base leads to a swaying or rolling motion in walking, which is seen in an exaggerated form in the waddling walk of a duck or a goose.

46. Machines. A machine is any device for so transforming or transferring energy as to enable man to employ the forces of nature for doing useful work. In all cases a certain amount of energy is applied *to* the machine and in return a certain amount of useful work is done *by* the machine. An ideally perfect machine would return all the energy applied to it. In actual practice a portion of the applied energy is always expended in overcoming friction, and is thus frittered away as heat. Notwithstanding this loss, machines are employed for the purpose of applying to better advantage the forces at our command.

Since energy is work or force × space, it is evident that with a given amount of energy the force may be *increased* exactly in the ratio that the distance through which it acts is *decreased.* Thus if P be the force applied *to* the machine, and s the distance through which it acts, W the force exerted *by* the machine, and s' the distance through which the machine works, then, if we neglect friction, we have

$$Ps = Ws' \tag{98}$$

or

$$\frac{W}{P} = \frac{s}{s'} \tag{99}$$

The ratio W/P between the force exerted and the force applied is called the *mechanical advantage* of the machine.

Newton described six elementary machines to which all others may be reduced. These are the *lever*, the *pulley*, the *inclined plane*, the *wheel and axle*, the *wedge* and the *screw*. These may be reduced to three, since the wheel and axle is but a modified lever, and the screw and the wedge are but modifications of the inclined plane. In all cases, however, the law of machines holds good, viz., *the force applied, multiplied into its acting distance, is equal to the force exerted multiplied into its acting distance.*

47. Simple Machines. *The lever.* The lever is a rigid bar turning freely about a fixed point F, called the fulcrum, which is to be considered as the center about which the moments of the forces P and W are to be taken.

For equilibrium these moments must be equal and opposite, or $\Sigma Fx = 0$

or
$$P \cdot AF = W \cdot FB \qquad (100)$$

from which the *mechanical advantage* is

$$\frac{W}{P} = \frac{AF}{FB} \qquad (101)$$

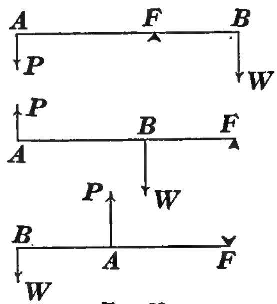


FIG. 23.

in each case. According to the relative position of the fulcrum and the points of application of the forces P and W, levers are classified as follows (Fig. 23):

(*a*) *Levers of the first class;* fulcrum between P and W. Examples, a crowbar, a pair of scissors, or a pair of forceps.

(*b*) *Levers of the second class;* W between the fulcrum and P. Examples, a pair of nutcrackers, a wheelbarrow, or a door swung by its outer edge.

(*c*) *Levers of the third class;* P between W and the fulcrum. Examples, a pair of sugar tongs, or a pair of tweezers; a door swung by its inner edge; the human forearm.

The pulley. A pulley is a small wheel turning freely about an axis, and having a grooved rim for the reception of

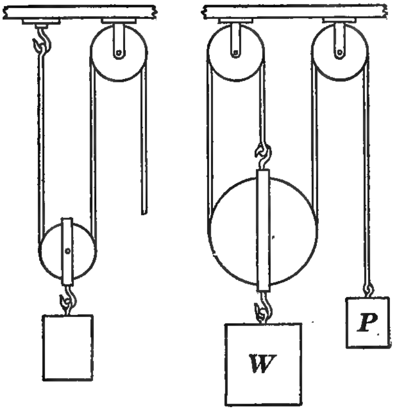


FIG. 24.

a cord. A fixed pulley serves simply to change the direction of the applied force. It is a lever with equal arms. A movable pulley (Fig. 24) permits of giving different velocities to P and W, and with a continuous cord its *mechanical advantage* is equal to the number of strands of cord supporting the weight.

The inclined plane. In the inclined plane we have a plane surface inclined at an angle θ to the horizon. Three cases are possible:

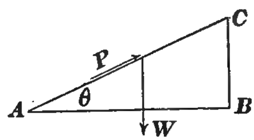


FIG. 25.

(*a*) *Force acts parallel to the plane.* In this case the force P (Fig. 25), tending to produce motion *up* the plane, is balanced by the component of the force of gravity *down* the plane, or for equilibrium

$$\Sigma F = 0$$

Hence we have

$$P = Mg(\cos 90° - \theta) = Mg \sin \theta \tag{102}$$

If we represent the length of the plane by l, the height by h, and the base by b, we have, on substituting the value of $\sin \theta$ in the above expression,

$$P = Mg\frac{h}{l} \tag{103}$$

or

$$Pl = Mgh \tag{104}$$

This shows that the work done in pushing the mass M up the plane is the same as that required to lift it through the vertical height h, whence the *mechanical advantage* is

$$\frac{Mg}{P}=\frac{l}{h} \tag{105}$$

(*b*) *Force acts parallel to the base of the plane.* In this case the force P (Fig. 26) furnishes a component up the plane which must be balanced by the component of the weight down the plane, or

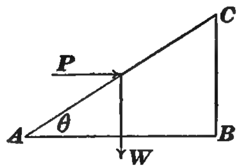


FIG. 26.

$$P\cos\theta=Mg\sin\theta \tag{106}$$

whence
$$P\frac{b}{l}=Mg\frac{h}{l} \tag{107}$$

or
$$P=Mg\frac{h}{b} \tag{108}$$

from which
$$Pb=Mgh \tag{109}$$

and again the work done in the two cases is equal. In this case we have for the *mechanical advantage*

$$\frac{Mg}{P}=\frac{b}{h} \tag{110}$$

(*c*) *Force acts at an angle ϕ with the plane.* In this case the equation of equilibrium is obtained as before by setting the components parallel to the plane equal to each other, or

$$P\cos\phi=Mg\sin\theta \tag{111}$$

The inclined plane permits of a small force being employed to do work by moving through a correspondingly greater distance. The winding curves of a railway as it makes its way up the mountain side illustrate the application of this principle.

48. Friction. Friction is the resistance offered to the motion of one portion of matter upon another. It is of the nature of a force. It seems to be due partly to the molecular attraction between the surfaces in contact, and partly to the nature and condition of those surfaces. Even the most highly polished surface is covered with minute scratches and projections which

oppose the motion of a similar surface over it. Friction always acts tangentially to the surfaces in contact and always opposes the motion.

Friction is of two kinds: (*a*) *static friction*, or the resistance offered to a force tending to start one body to slide uniformly upon another; (*b*) *kinetic friction*, or the force necessary to keep the body sliding in uniform motion on a horizontal surface. Static friction is always greater than kinetic friction and varies from zero up to a maximum value, but is never more than that required to prevent motion.

The following facts concerning friction seem well established by experiment: (*a*) The friction depends upon the nature of the substances and the condition of the surfaces, but is independent of the area of the surfaces in contact.

(*b*) The friction is independent of the speed of the motion unless it be very small, when it rises to the maximum or static friction as the speed falls to zero.

(*c*) The friction is proportional to the force with which the surfaces are pressed together that is, it is that force times some factor which is always less than unity. This factor is called the *coëfficient of friction.*

49. The Balance. The balance is an instrument for the comparison of masses. It is a lever of the first class, having equal arms. In its simplest form it consists of a light, strong beam, supported at its center by a knife-edge resting upon an agate plate, and carrying upon knife-edges at its ends two pans for the reception of the masses to be compared. When disturbed the system oscillates about its position of equilibrium, to which it finally returns. When masses are placed in the pans, it is evident that the original position of equilibrium will be resumed only when the moments of the forces of gravity acting upon these masses are equal. If we assume the arms of the balance to be equal, we may set the masses equal to each other when their moments have been shown to be equal; that is, when the balance resumes its original position of equilibrium.

In case the arms of the balance are not of equal length, the true weight of the body may yet be determined by the method

of double weighing. Let the arms of the balance be l and l_1, and let the true weight of the body be W. Suppose the body when hung from the arm l, to be balanced by a weight W_1 on l_1, and when hung from the arm l_1, to be balanced by a weight W_2 on l. Then by the principle of moments we have for the first case

$$Wl = W_1 l_1 \tag{112}$$

and for the second

$$W_2 l = W l_1 \tag{113}$$

Dividing the first equation by the second, we have

$$\frac{W}{W_2} = \frac{W_1}{W} \tag{114}$$

or

$$W^2 = W_1 W_2 \tag{115}$$

whence

$$W = \sqrt{W_1 W_2} \tag{116}$$

From this it appears that the true weight of a body may be found upon a balance of unequal arms by weighing first in one pan and then in the other and taking the square root of the product. If W_1 and W_2 are not greatly different from each other, the arithmetical mean of the two weights is usually sufficient unless extreme accuracy is required.

***50. Sensibility of the Balance.** The delicacy of the balance is determined by its sensibility. This depends upon a number of considerations and is expressed in terms of the angular displacement of the beam for a small difference of weight p, usually one milligram, in the pans.

The sensibility of a balance is determined as follows: Let the points A, B, C (Fig. 27), represent the knife-edges of the two pans and the beam, and suppose these points to lie in a horizontal line. Let the center of gravity of the beam and pointer of the balance be at G, distant r

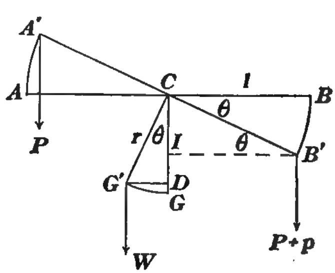


FIG. 27.

from the center of suspension C, and let the weight of this system be W. For two equal weights P and P', placed in the pans, the system remains in equilibrium, provided the arms be of equal length. If an additional small weight p be placed in the right-hand pan, the system is displaced through the angle θ, and finally comes to rest in the position $A'CB'$. In this position the restoring moment due to W, and the disturbing moment due to p, must be equal and opposite. Then

$$W \times G'D = p \times B'I \tag{117}$$

Substituting for $B'I$ and $G'D$ their values in terms of θ, we have

$$Wr \sin \theta = pl \cos \theta \tag{118}$$

whence

$$\frac{\tan \theta}{p} = \frac{l}{Wr} \tag{119}$$

The quantity $\frac{\tan \theta}{p}$ is termed *the sensibility of the balance*[1] and is seen to vary directly as the length of the beam, and inversely as the weight of the beam and pointer, and as the distance of the center of gravity of the system, from the point of support.

In order, therefore, that a balance should be sensitive, the beam should be long and light, and the distance r, between the center of gravity of the system and the center of support, should be small. Practical considerations, however, set a limit to each of these theoretical suggestions. The beam must neither be so long nor so light as to render it liable to bend under the load it is to carry, else the sensibility would be diminished; also if the distance r be made too small, the system becomes unstable and the time of vibration becomes inconveniently long.

51. Moment of Inertia. It is now necessary to investigate the expression for the kinetic energy of a rotating body. Suppose a rigid body to rotate uniformly about a line through its

[1] *For specific directions in the use of the balance, see Manual, Exercises 12 and 13.*

center of gravity, O (Fig. 28), and let us call this line the axis of rotation. Also let ω denote the angular velocity. Then two particles of masses m_1 and m_2, situated at distances r_1 and r_2 from this axis, describe circles of circumferences $2\pi r_1$ and $2\pi r_2$ respectively, *in the same time*, T. The linear velocities v_1 and v_2 of the particles are consequently

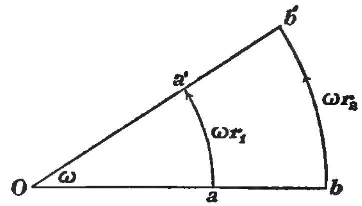


FIG. 28.

$$\frac{2\pi r_1}{T} \text{ and } \frac{2\pi r_2}{T}$$

But since

$$\frac{2\pi}{T} = \omega$$

we have

$$\begin{aligned} v_1 &= \omega r_1 \\ v_2 &= \omega r_2 \end{aligned} \tag{120}$$

or, in general, the *linear velocity* of any point revolving about an axis is equal to the angular velocity times the radius; or,

$$v = \omega r \tag{121}$$

Again, the kinetic energies of the two particles m_1 and m_2 will be

$$\frac{m_1 v_1^2}{2} = \frac{m_1 \omega^2 r_1^2}{2}$$

and (122)

$$\frac{m_2 v_2^2}{2} = \frac{m_2 \omega^2 r_2^2}{2}$$

If we suppose the rotating body to be made up of an indefinite number of particles, then the kinetic energy of the body will be the sum of the kinetic energies of all its particles, or

$$K.E. = \frac{m_1 \omega^2 r_1^2}{2} + \frac{m_2 \omega^2 r_2^2}{2} + \frac{m_3 \omega^2 r_3^2}{2} + \cdots + \frac{m_n \omega^2 r_n^2}{2} \tag{123}$$

or, since ω is constant for the entire body,

$$K.E. = \frac{\omega^2}{2}\Sigma mr^2 \tag{124}$$

The expression Σmr^2 is called the *moment of inertia*[1] of the body, and is designated by the letter I. It denotes the sum of the products, obtained by multiplying the mass of each individual particle by the square of its distance from the axis of rotation. This quantity, Σmr^2, has a perfectly definite, positive, numerical value for a definite body, rotating about a definite axis. It will be observed that I depends, not only upon the *mass* of the body but much more upon the *manner in which that mass is distributed with regard to the axis of rotation.* For every body it is possible to find a radius K, such that the mass of the body M multiplied by the square of this radius is equal to the moment of inertia about the axis of rotation, or

$$MK^2 = \Sigma mr^2 \tag{125}$$

Such a radius is called the *radius of gyration.*

The dimensions of a moment of inertia are $[ML^2]$ and the unit is gram centimeter square.

For convenience of reference the following table of moments of inertia is appended, where M denotes the mass of the body and I the required moment of inertia.

Moments of Inertia (Formulae)

Uniform thin rod, axis through middle, length $= l$,

$$I = M\frac{l^2}{12} \tag{126}$$

Uniform thin rod, axis at one end, length $= l$,

$$I = M\frac{l^2}{3} \tag{127}$$

Rectangular lamina, axis through center and parallel to one side, length of sides a and b, side b bisected,

[1] *For experimental determination of moments of inertia, see Manual, Exercises 25 and 26.*

$$I = M\frac{b^2}{12} \tag{128}$$

Rectangular lamina, axis through center and perpendicular to the plane, sides a and b,

$$I = M\frac{a^2 + b^2}{12} \tag{129}$$

Circular plate, axis through center perpendicular to the plate, radius $= r$,

$$I = M\frac{r^2}{2} \tag{130}$$

Circular plate, axis any diameter, radius $= r$,

$$I = M\frac{r^2}{4} \tag{131}$$

Circular ring, axis through center perpendicular to plane of ring, outer radius $= R$, inner radius $= r$,

$$I = M\frac{(R^2 + r^2)}{2} \tag{132}$$

Circular ring, axis any diameter, radii as before,

$$I = M\frac{(R^2 + r^2)}{4} \tag{133}$$

Moment of inertia about an axis parallel to an axis through center of gravity of body, and distant from same by a distance a, moment of inertia about axis through center of gravity $= I_0$,

$$I = I_0 + Ma^2 \tag{134}$$

52. Moment of Inertia and Angular Acceleration. If α be the angular acceleration, then the linear acceleration

$$a = \frac{v - v_0}{t} = \frac{\omega r - \omega_0 r}{t} = \alpha r \tag{135}$$

or the linear acceleration is equal to the angular acceleration times the radius; that is,

$$a = \alpha r \tag{136}$$

If now we consider a particle of mass m, at a distance r from the axis of rotation, then the force necessary to give it a linear acceleration a is

$$F = ma = m\alpha r \tag{137}$$

and the moment of the force is $m\alpha r \times r$, since the force acts tangentially to the circle described by m, or the moment *for a single particle* is $m\alpha r^2$.

The moment of the force necessary to give the entire body the same acceleration is $\Sigma m\alpha r^2$, or $\alpha\Sigma mr^2$; that is,

$$\textit{Moment} = \textit{Torque} = \mathscr{T} = I\alpha \tag{138}$$

53. Kinetic Energy of Rotation. We have now derived two distinct formulae for kinetic energy.

(*a*) $$\textit{K.E. of translation} = \tfrac{1}{2}Mv^2 \tag{139}$$

(*b*) $$\textit{K.E. of rotation} = \tfrac{1}{2}I\omega^2 \tag{140}$$

On comparing these formulae for kinetic energy, we note from their symmetry that the angular velocity ω of a rotating body corresponds to the linear velocity v of a body undergoing translation; *also that the moment of inertia I of the rotating body corresponds to the mass M of the body undergoing translation.* This relation is most important, and finds constant application in mechanics. The flywheel of an engine acts as a reservoir of kinetic energy, not only on account of the mass, but still more on account of the distribution of its mass, since this is largely in the rim, thus making its moment of inertia as large as possible.

Again, it should be observed that a system may possess kinetic energy both of translation and of rotation at the same time. Thus the wheel of a bicycle when in motion possesses kinetic energy due to its motion of translation equal to $\frac{1}{2}M\bar{v}^2$, and also due to its rotation amounting to $\frac{1}{2}I\omega^2$. This fact has been ingeniously employed in the manufacture of steel shells for modern, long-distance, rifled cannon. The flying cylindrical shell possesses both energy of translation and energy of

spin around its longer axis, due to the twisting motion imparted to it by the rifling of the gun. The striking end of the shell is now made of the hardest steel, and furnished with a screw point, so that the instant it strikes a target, the kinetic energy of rotation causes it to bore in like an auger, thus increasing its destructive action many fold.

The following table shows at a glance the striking analogy between the formulae relating to *linear* and *angular* motion:

TABLE II

TRANSLATION	ROTATION	TRANSLATION	ROTATION
s	θ	$s = v_0t + \frac{1}{2}at^2$	$\theta = \omega_0t + \frac{1}{2}\alpha t^2$
v	ω	$F = Ma$	$\mathcal{T} = I\alpha$
a	α	$W = Fs$	$W = \mathcal{T}\theta$
F	$\mathcal{T}$	$\mathcal{P} = Fv$	$\mathcal{P} = \mathcal{T}\omega$
M	I	$K.E. = \frac{1}{2}mv^2$	$K.E. = \frac{1}{2}I\omega^2$

54. Ideal Simple Pendulum. The ideal simple pendulum may be defined as a heavy material particle suspended by a weightless thread. As an approximation to this ideal we may use a lead ball supported by a fine silk cord. It is required to find an expression for the time of vibration of this pendulum. Let the pendulum of mass m be supported by a cord AB, of length l (Fig. 29). When drawn aside to the position B, it forms an angle θ, or BAN, with the vertical. The component of the weight mg, which is effective in producing motion, is $mg \cos (90° - \theta)$ or $mg \sin \theta$. Let NB equal x, the displacement, and let a_x be the acceleration toward the position of rest; then ma_x is the force urging the body toward N, or

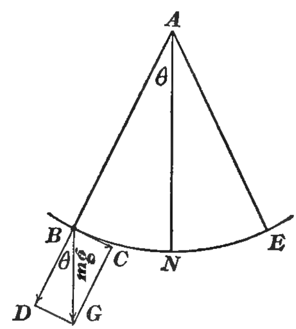


FIG. 29.

$$ma_x = - mg \sin \theta \tag{141}$$

whence $$a_x = -g \sin\theta \tag{142}$$

also $$\theta = \frac{x}{l} \tag{143}$$

If θ be small, we may set

$$\sin\theta = \theta = \tan\theta$$

therefore $$a_x = -\frac{g}{l} \cdot x \tag{144}$$

Now since g and l are constants, it is clear that within the limits where we may write

$$\sin\theta = \theta = \tan\theta$$

the *acceleration is proportional to the displacement*, and the motion of the pendulum will be *simple harmonic motion.* But from equations (69) we have

$$a_x = -\frac{4\pi^2}{T^2} \cdot x$$

therefore $$\frac{4\pi^2 \cdot x}{T^2} = \frac{g}{l} \cdot x \tag{145}$$

or $$T = 2\pi\sqrt{\frac{l}{g}} \tag{146}$$

This is the *period*, or the time of a *complete vibration*;[1] for a half vibration

$$T_1 = \pi\sqrt{\frac{l}{g}} \tag{147}$$

To find the length of the seconds pendulum, where

$$g = 980\frac{\text{cm}}{\text{sec}^2}, \text{ set } T_1 = 1, \text{ or } 1 = \pi\sqrt{\frac{l}{980}}$$

from which $$l = \frac{980}{\pi^2} = 99.3 \text{ cm}$$

55. Compound or Physical Pendulum. Any body suspended so as to swing freely about an axis forms a compound or physi-

[1] *For experimental verification of this formula, see Manual, Exercise 23*

cal pendulum. In such a pendulum the weight of the body may be considered as a force acting on its center of gravity. Suppose the heavy body (Fig. 30) be drawn aside through an angle θ, then the moment of the restoring force is $Mgh \sin \theta$, where h is the distance from the point of suspension to the center of gravity of the pendulum. This same moment tends to produce an angular acceleration α, and we have seen (Art. 52) that such a moment is equal to $\alpha \Sigma mr^2$, hence equating these two values for the moment of the force Mg upon the pendulum, we have

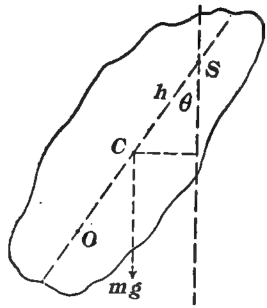


FIG. 30.

$$\alpha \Sigma mr^2 = Mgh \sin \theta \text{ and } \frac{\Sigma mr^2}{Mh} = \frac{g \sin \theta}{\alpha} \tag{148}$$

Now let us consider a simple pendulum whose mass is equal to the mass M of the physical pendulum and whose length is l. If it be displaced through an angle θ the torque acting on it is

$$Mgl \sin \theta = Ml^2 \alpha \text{ and } l = \frac{g \sin \theta}{\alpha} \tag{149}$$

If α be the same as for the physical pendulum, l is evidently the length of an equivalent simple pendulum, which will vibrate in the same time as the body. If we let

$$l = \frac{g \sin \theta}{\alpha} = \frac{\Sigma mr^2}{Mh} = \frac{I}{Mh} \tag{150}$$

equal the length of this simple pendulum, we may write the expression for the time of vibration of the body at once.

$$T = 2\pi \sqrt{\frac{l}{g}} = 2\pi \sqrt{\frac{I}{Mhg}} \tag{151}$$

This is a general expression for the time of vibration of any heavy body when disturbed, about its position of equilibrium.

The length of a compound pendulum may be determined experimentally by suspending near it a simple pendulum and

altering the length of the latter until the two vibrate in the same time. The length of the simple pendulum so found is

$$l = \frac{\Sigma mr^2}{Mh} = \frac{I}{Mh}$$

If now we lay off this length upon the compound pendulum, measuring *from* the center of suspension *through* the center of gravity, we find the *center of oscillation* O; that is, the center of oscillation is distant from the center of suspension by the length l, as defined by equation (150), and the pendulum behaves as if all the mass were located at this point.

It may be also shown that the center of oscillation and the center of suspension are interchangeable; or, in other words, for every point of suspension S there exists a conjugate point O, such that the pendulum will vibrate in the same time, whether suspended from S or from O.

The center of oscillation is also termed the *center of percussion*, since, if the pendulum be struck a blow at this point, it will move forward in its arc, with no jerk or jar upon the point S, as if the whole mass had received the blow. At any other point the effect of a blow would be to start the pendulum to swinging, and also to jerk it from its support owing to the tendency to rotate about some axis other than the point of support. This is well illustrated in the case of a baseball bat, which is to be considered as a compound pendulum. If the ball be received upon the bat at the center of percussion, no jarring of the hands is experienced by the striker. If the ball strike the bat at any other point, either too near the end or too near the hands, the jarring effect is very painful, and the bat is usually broken as well.

Problems

1. A weight of 350 kilos is carried by a bar, AB, 4.7 m long, supported at the ends. If the distance of the weight from the end A be 2 m, what is the weight borne by each support? *Ans.* 201.1; 148.9 kilos.

2. A uniform lever 8.2 ft long weighs 15 lb. If 26 lb be hung from one end and the bar be sustained by a support 2.6 ft from this end, what weight must be hung from the other end for the system to be in equilibrium? *Ans.* 8.05 lb.

3. It is found that a lever cut from a uniform bar weighing 4.2 lb to the foot balances at a distance of 2.3 ft from one end when weighted at this end with 5.4 lb. What is the length of the bar? *Ans.* 5.65 ft.

4. When a lever, AB, is supported at its center of gravity, it is found that a weight, W, hung at A will balance 2.5 lb at B; but when W is hung at B, it requires a weight of 19 lb at A to keep it in equilibrium. What is the weight W? *Ans.* 6.89 lb.

5. Find the magnitude and point of application of the resultant of two parallel forces, in the same direction, equal, respectively, to 25 and 42, when their lines of action are 3.4 ft apart. *Ans.* 67 applied 2.13 ft from force 25.

6. A meter scale (of which the mass may be neglected) has suspended from the 10 cm mark, and from the 90 cm mark masses of 2 and 9 kilos respectively. Where must the scale be supported for equilibrium?

Ans. 75.45 cm.

7. A lever is 20 cm long and its mass 15 g. Where must the fulcrum be placed in order that a mass of 10 g at one end may just balance a mass of 16 g at the other end? *Ans.* 8.54 cm from the 16 g mass.

8. A uniform, thin rod one meter long, whose mass may be neglected, carries four masses of 4, 7, 10 and 22 g. The 4 g mass is at one end and the 22 g at the other, while the 7 and 10 g masses are 55 and 75 cm from the 4 g mass. Find the center of inertia of the system.

Ans. 77.56 cm from 4 g mass.

9. From a square plate of uniform thickness, and area a^2 cm², one triangle formed by the intersection of the diagonals has been removed. Find the center of inertia of the remainder. *Ans.* $\frac{a}{9}$ cm from center on a line bisecting triangle opposite triangle removed.

10. Masses of 7, 9 and 11 kilos are placed at the points A, B and C of a right triangle, right angled at C, of which the base AC is 10 cm and the altitude BC is 15 cm, and whose mass may be neglected. Find the center of inertia of the system.

Ans. 5 cm from AC and 2.593 cm from BC.

11. A wheel of mass 100 lb is attached to end of a straight uniform axle 4 ft long and of mass 50 lb. Find the center of inertia of the system.

Ans. 2/3 ft from wheel.

12. In a system of one fixed and one movable pulley (Fig. 24), what weight can be lifted by a force of 20 lb weight, neglecting friction and mass of pulley? *Ans.* 40 lb.

13. In a system of two fixed and one movable pulley (Fig. 24), what mass at end of cord will equilibrate 270 kilos attached to movable pulley, neglecting friction and mass of pulley as before? *Ans.* 90 kilos.

14. What force acting horizontally can keep a mass of 16 kilos at rest on a smooth inclined plane, whose height is 3 m and base 4 m? *Ans.* 12 kilos.

15. Find the force normal to the plane mentioned in problem 14, if the mass be kept at rest by the horizontal force. *Ans.* 20 kilos.

16. A body of mass 25 kilos rests upon an inclined plane, which makes an angle of 30° with the horizon. Find the force necessary to produce equilibrium, (*a*) when the force acts parallel to plane, (*b*) when parallel to base. *Ans.* (*a*) 12.5 kilos.
(*b*) 14.43 kilos.

17. A boy who can exert a force equal to 50 lb weight, wishes to roll a barrel of mass 200 lb into a window 4 ft high. How long a plank is required? *Ans.* 16 ft.

18. If a mass weigh 29.62 g, in one pan of a balance and 28.71 g in the other, what is its true mass? *Ans.* 29.162 grams.

19. If a mass of 625 g moving with a velocity of 786 cm per second meet a mass of 164 g moving in the opposite direction with same speed, what will be their velocity after impact, supposing the bodies to be inelastic? *Ans.* 459.25 cm per sec.

20. A grindstone of radius 20 cm and density 2.2 g per cubic centimeter, is 10 cm thick. Calculate its moment of inertia when rotated about an axis through the center. *Ans.* 5,529,216 g-cm^2.

21. If this stone rotate 5 times in 4 min, compute its kinetic energy. *Ans.* 47,371 ergs.

22. Compare the moment of inertia of a meter rod of mass 150 g, when rotated about an axis, (*a*) through one end, and (*b*) through the center, both axes being normal to the length of the rod. *Ans.* As 4 to 1.

23. A board 2 cm thick and 8 cm long by 5 cm wide, of density 0.75 g per cubic centimeter, rotates about an axis, through its center of inertia and normal to its plane. Compute its moment of inertia. *Ans.* 445 g-cm^2.

24. A grindstone of mass 25 kilos and radius 40 cm rolls without sliding down an inclined plane 200 cm long and 50 cm high. Compare its kinetic energies of rotation and translation. *Ans.* As 1 to 2.

25. Find the time of a single vibration of a simple pendulum 20 cm long, where g equals 980 cm per second per second. *Ans.* $T = 0.4488$ sec.

26. What would be the value of g, if the period of a simple pendulum 97.31 cm in length were 1.975 sec? *Ans.* 984.88 cm per sec^2.

27. A pendulum loses 20 sec per day, where g equals 980.3 cm per second per second. Find its length. *Ans.* 99.4 cm.

28. A compound pendulum is composed of a wire 60 cm long, whose weight may be neglected, suspended from one end, and carrying masses of 3, 5 and 7 kilos at distances of 20, 40 and 60 cm from the point of support. Compute the length of the equivalent pendulum and time of double vibration. *Ans.* (*a*) 50.59 cm.
(*b*) 1.43 sec.

29. A thin hoop of radius 30 cm and mass 100 g hangs over a fine nail and vibrates in its own plane. Find its time of vibration. *Ans.* 1.556 sec.

30. A hollow cylinder rolls down a perfectly rough inclined plane in 10 min. Find the time a uniform solid cylinder would take to roll down the same plane. *Ans.* $5\sqrt{3}$ minutes.

31. The masses in an Atwood's machine are 355 and 345 g. In the *second* second, starting from rest, the heavier mass descends through 20 cm. What value does this give for g? *Ans.* 933.3 cm per sec^2.

CHAPTER VI

ELASTICITY

56. Stress and Strain. Elasticity is that property by virtue of which matter resists the action of a force tending to change its shape or bulk, and which causes it to resume its original shape or bulk after the force is removed. If a body possess elasticity of shape, it is called *a solid;* if it possess no elasticity of shape, it is called *a fluid.* If a body have its size or shape changed under the action of a force, it is said to have been *strained.* The change in size or shape is called *a strain*, and is measured by the *relative change* so produced. Thus, if a mass of gas of volume V have its volume changed by a small amount dV, then the strain is $\frac{dV}{V}$; or if a solid of length L have its length changed by dL, then the strain is $\frac{dL}{L}$.

The external force applied to produce a strain is called the *load*, and, as a result of the strain produced, there arises in the strained medium a system of *resisting forces* which tend to produce equilibrium. This reaction of the medium against further deformation is called *a stress.* The stress at any point in a medium is measured by the *force per unit area.* After equilibrium ensues the load and the resisting force are equal. Hence stress may be evaluated in terms of the *applied force per unit area.* In the case of a compressed gas this stress is expressed as the *pressure* exerted by the gas in $\frac{\text{dynes}}{\text{cm}^2}$, or $\frac{\text{lb}}{\text{in}^2}$.

Fluids possess *perfect elasticity of bulk;* that is, they return exactly to their original bulk on removal of the compressing load. For every solid there is a limiting distortion beyond which the body when freed from the distorting load no longer

completely regains its former shape. This distortion is called *the limit of elasticity*. In engineering practice the *elastic limit* is expressed in terms of the *stress* producing this limiting distortion.

All solids do not recover their initial shape with equal promptness. In some cases this return is much retarded, especially after repeated or long-continued distortion. This is commonly termed *elastic after-effect*, and is quite noticeable in metals.

57. Hooke's Law. When an elastic body is distorted within its limit of elasticity, the stress called out by the distortion, tending to restore the body to its original condition, is *proportional to the distortion.* This is known as *Hooke's law*, and as originally stated, "*ut tensio sic vis*," expresses the proportionality between the stress and the strain. The applications of this law are very numerous, including every form of elastic reaction against strains produced by external mechanical agencies.

As will be readily seen, the application of Hooke's law involves the fundamental assumption that the force, and therefore the *acceleration*, is proportional *to the distortion.* If the original position of the system be taken as the position of rest, we have at once the characteristic condition of simple harmonic motion. Such motions are seen in the vibrations of elastic springs, of the air particles in the production of sound, in the varying charges in an electric condenser, and in compressional and torsional waves in solids.

From Hooke's law we see that in all cases where a body is distorted within its limits of elasticity, we have a relation of the form

$$stress = c \times strain \tag{152}$$

where c is a *constant* which is independent of either the stress or the strain, but which is characteristic of the substance undergoing distortion. It is called *a coefficient of elasticity.*

58. Coefficients of Elasticity. In general, twenty-one coefficients would be necessary to express completely the elastic nature

of a body. If, however, the body be isotropic, these twenty-one coefficients reduce to two: the *coefficient of volume elasticity e, and the coefficient of rigidity n.* The general expression for these coefficients is therefore the quotient arising from *dividing the stress by the strain.*

In the coefficient of volume elasticity e we have the stress, or applied pressure p, divided by the compression produced, where *compression* denotes the *change in volume v, divided by the original volume V*, or

$$e = \frac{p}{\frac{v}{V}} = \frac{p}{v} V \tag{153}$$

Since all changes in volume are conceived as being very small, if we assume a volume of gas V, to be subjected to a change in pressure dp, producing a corresponding change in volume dV, then

$$e = \frac{dp}{dV} V \tag{154}$$

It should be observed that the expression for the strain $\frac{dV}{V}$ denotes a *dilatation if positive*, and a *compression if negative*. The coefficient e, however, has reference simply to the numerical magnitude of the ratio, $\frac{dp}{\frac{dV}{V}}$, and is therefore independent of the sign.

Besides the elasticity of volume, solids have, as we have seen, elasticity of shape as well. If a solid be so distorted that its shape alone is changed, it is said to have undergone a *shear*. Thus if we conceive all the particles in a certain plane in a body to be fixed, and all the remaining particles to move in planes parallel to this plane, and by amounts proportional to their distances from this plane, the resulting distortion is called *a shear*. The stress arising from such a distortion is called a *shearing stress*, and the coefficient of rigidity n is the quotient obtained by dividing the shearing stress by the shearing strain.

59. Young's Modulus. If a load

$$F = mg$$

be suspended from a wire of uniform cross section a and length L, the wire lengthens by an amount l, and a strain is produced in the wire, equal to $\frac{l}{L}$. The stress due to this distortion resists the action of the stretching force, and equilibrium is established when the force due to the stress becomes equal to the load; or if

$$F = stress \times a = mg.$$

The coefficient of linear stretch M is given by the equation

$$M = \frac{stress}{strain} = \frac{\frac{mg}{a}}{\frac{l}{L}} = \frac{mgL}{al}. \tag{155}$$

This coefficient is called Young's modulus,[1] and its unit is the dyne per cm². Under otherwise equal conditions the elongation of a wire will be smaller, the larger Young's modulus is for the substance of which the wire is made.

TABLE III

COEFFICIENTS OF ELASTICITY

SUBSTANCE	YOUNG'S MODULUS	COEFF. OF VOLUME ELASTICITY
Brass	$11 \times 10^{11} \frac{\text{dynes}}{\text{cm}^2}$	$10.5 \times 10^{11} \frac{\text{dynes}}{\text{cm}^2}$
Steel	$22 \times 10^{11} \frac{\text{dynes}}{\text{cm}^2}$	$18.5 \times 10^{11} \frac{\text{dynes}}{\text{cm}^2}$
Cast Iron	$13.5 \times 10^{11} \frac{\text{dynes}}{\text{cm}^2}$	$9.5 \times 10^{11} \frac{\text{dynes}}{\text{cm}^2}$
Copper	$12 \times 10^{11} \frac{\text{dynes}}{\text{cm}^2}$	$16.5 \times 10^{11} \frac{\text{dynes}}{\text{cm}^2}$

60. Three States of Matter. Matter presents itself in one of three states: *solid*, *liquid* or *gaseous*. These states are not separated by sharp lines of demarcation, but shade almost insen-

[1] *For experimental determination of Young's modulus, see Manual, Exercises 18 and 20.*

sibly into each other, the particular state which the body assumes being dependent upon the temperature and pressure to which it is subjected.

A solid is a body which resists any force tending to change either its shape or its volume; in other words, it possesses elasticity both of shape and of bulk. It is characterized by small mobility of its molecules, and by the possession of surfaces of distinct outline on all sides.

A liquid is a body which has no elasticity of shape, or which offers no resistance to a shearing force; it is characterized by considerable mobility of its molecules, by a distinct, free, upper surface, usually of a meniscus shape when confined in a tube, and by the existence in that free surface of a specific tension, not found elsewhere in the body.

A gas is a body whose parts are held together only by the action of external force, and which offers resistance to the *decrease* but none to the *increase* of its volume. It is characterized by almost perfect mobility of its molecules. It has no upper surface and no surface tension, but tends to fill all available space.

Liquids and gases when subjected to shearing forces are unable to remain in equilibrium, and their parts move into new positions, or tend to *flow;* they are therefore classed together under the general term *fluids.* Fluids have perfect elasticity of volume but no elasticity of shape.

61. Intermediary Qualities. As has been said, the three states of matter shade gradually into each other, although the changes are more abrupt in some cases than in others. Thus while most substances are perfectly elastic under hydrostatic pressure, no substance has yet been found to be perfectly elastic under shearing forces producing finite strains. Owing to the varying behavior of substances under stress, a number of intermediary qualities are manifested, by virtue of which certain characteristics of the fluid state are seen to exist even in solids.

A body which breaks upon the sudden application of slight force is said to be *brittle.* Glass, gems and hardened steel are examples of brittle substances.

A *plastic* substance is one which has its form permanently altered by a force exceeding a certain value. Plastic solids under pressure behave as liquids, and flow with greater or less freedom. Thus cold lead under enormous pressure is made to flow through a die, producing lead pipe.

Plastic bodies may be beaten into various forms, rolled into sheets, or drawn into wire. This is exemplified in the case of copper and platinum and in still higher degree in the case of gold. Substances which may be beaten into thin sheets are called *malleable;* those that may be drawn into wire are said to be *ductile.* Certain substances are plastic at high temperatures and brittle at ordinary temperatures. Thus glass and quartz are extremely brittle at ordinary temperatures, yet when fused they may be drawn into threads of exceeding fineness. Quartz especially, when fused, may be drawn into fibers so fine as to be invisible to the naked eye, yet capable of supporting relatively large masses, and possessing almost entire freedom from elastic after-effect.

Certain organic substances rival the metals in ductility. Thus sugar and pitch at certain temperatures may be drawn into threads so fine as to float on the air, while the spider's thread is formed in some cases of a thousand separate strands united into one.

62. Viscosity. While it is assumed that an ideal liquid would offer no resistance to a shearing stress, still it is to be noted that all known liquids do manifest a certain degree of resistance to the motion of their parts over each other. This *internal friction*, or *resistance to flow*, is called *viscosity.* A liquid that flows readily as alcohol or ether is termed a *limpid* liquid as opposed to a *viscous* liquid like honey or Canada balsam.

To illustrate the viscosity of liquids, support horizontally a glass disk 10 cm in diameter, by a stiff wire from its center, and lower it into a flat beaker containing several centimeters depth of glycerine, bringing the disk at least 5 cm below the surface of the liquid. Place small bits of paper or cork upon the glycerine and rotate the disk by twisting the supporting

wire. The surface is soon seen to be in rotation, showing that the rotation of the disk sets up a shearing stress which is transmitted through the liquid to the surface layers and sets them in motion also. A smooth metal disk mounted horizontally upon a whirling table and set in rapid rotation will drag along with it the adjacent layers of air, as may be shown by lowering over it a similar disk suspended by a fine cord from its center. The viscosity of the air will soon suffice to set the upper disk in rotation even at a distance of one or two centimeters.

It thus appears that a column of glycerine or even a column of air is able, by virtue of the internal friction between its molecules, to transmit a shearing stress from one end of the column to the other. In the case of a perfectly rigid column such transmission would be instantaneous, while in the case of fluid columns there is a continual breaking down under the stress and the transmission is much slower and far less perfect.

This continued breaking down or yielding of the parts of a body under the action of a shearing stress is seen also in the case of bodies usually classed as solids. Thus shoemaker's wax, while brittle to a sudden blow, will creep slowly down an inclined plane under the action of gravity; long glass tubes, laid horizontally upon supports at the ends, sag down in the middle, due to their own weight; sandstone pillars under enormous weights have slowly bent without breaking. All these are examples of viscous substances.

***63. Coefficient of Viscosity. Poiseuille's Law.** Internal friction likewise determines the velocity with which a liquid under pressure will flow through a tube of narrow bore. Since the layer of liquid in contact with the tube remains practically at rest, the velocity of flow increases from without inward, becoming a maximum at the center. With a definite difference of pressure between the two ends of the tube, the liquid assumes a steady state of flow, and the moving mass may be considered as made up of a number of concentric liquid tubes slipping over each other under pressure and retarded by molecular friction. With a difference of pressure p, between the two ends of the

tube, the quantity Q, discharged through a tube of length l, and radius r, in time t, is given by the equation

$$Q = \frac{\pi p r^4}{8 l \eta} \cdot t \tag{156}$$

where η is called the *coefficient of viscosity.*

This relation, first determined experimentally by Poiseuille, a French physician, and subsequently derived theoretically by others, is known as *Poiseuille's law.* It has been found to hold very accurately for small velocities of flow, but fails when the velocity exceeds a certain value depending upon the size of the tube and the nature of the liquid. Poiseuille's law is of importance in questions relating to the capillary circulation of the blood.

The coefficient of viscosity η is defined as numerically equal to the tangential force exerted upon unit area of either of two horizontal planes placed at unit distance apart, in the viscous substance, one of the planes being fixed and the other moving with unit velocity. This coefficient may be measured in several other ways, as by determining the resistance encountered by a solid in moving through the viscous liquid.

The value of η varies greatly for different substances. For water at 15° C its value is 0.0115; for glycerine it is about 300 times as much; for ether one third as much; and for air one fifty-fifth as much. The viscosity of liquids usually diminishes with increase of temperature. Hot water is less viscous than cold; sirups and saline solutions are limpid when hot and viscous when cold. In gases, on the other hand, the viscosity *increases* as the temperature increases.

MECHANICS OF FLUIDS

CHAPTER VII

FLUIDS AT REST

64. Fluid Pressure. A perfect fluid is a body which would offer no resistance to a shearing stress. This, as we have seen, is a purely ideal case, but if we consider the distortions in a fluid mass as vanishingly small, even the slight resistance of the fluid to change of form arising from viscosity becomes zero, and we find that *fluids at rest behave as perfect fluids.* Since in a perfect fluid there is no viscosity, and hence no statical friction, it follows that the *pressure* exerted by the fluid on the containing vessel must all be *normal* to the *surface* of the vessel, because any pressure other than a normal pressure could be resolved into a normal and a tangential component. The normal component would produce pressure on the side of the vessel; the tangential component would produce motion in the fluid unless opposed by statical friction. But there is no statical friction, and since the fluid *is at rest*, there can be *no tangential component.* Therefore the fluid pressure on the side of a vessel is normal to the side of the vessel.

Again, it follows from the absence of statical friction in a fluid that any force exerted upon any surface of a confined fluid must be transmitted undiminished to every equal area in contact with the fluid. For consider two cylinders connected with a closed vessel (Fig. 31), each cylinder being stopped by a piston of area a, and the whole filled with a fluid. If now one piston be forced in through a distance b, then the second piston must be driven out an equal distance, if no compression of the fluid is to ensue. The work done by the first piston is

Fb, where F is the force applied. Since no work is done in compressing the liquid, the work done on the second piston must be $F'b$. Whence

$$Fb = F'b \tag{157}$$

or

$$F = F' \tag{158}$$

Even if the pistons have unequal areas, a and a', the pressure is still transmitted undiminished, since in this case we have

$$Pab = P'a'b' \tag{159}$$

and ab, the volume of liquid displaced on one side, is equal to $a'b'$, the volume displaced on the other side, if no compression takes place. Whence

$$P = P' \tag{160}$$

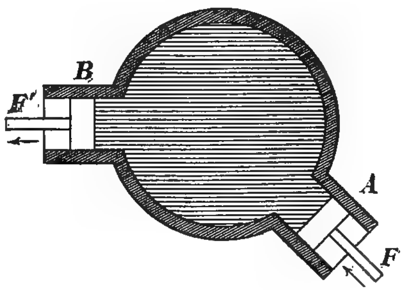


FIG. 31.

65. Pressure at any Point in a Fluid. Since pressure in a fluid is transmitted undiminished in all directions, it follows that the pressure at any point is the same in all directions. For suppose a body of liquid to be in equilibrium, and conceive a cylinder of the fluid of very small diameter to become solid without change of density. Then unless the pressures upon the ends of the cylinder are equal and oppositely directed, motion will ensue. But the liquid is at rest; hence there is no difference in the pressures upon the two ends. Now since the cylinder may be rotated in a horizontal plane, it follows that at any point in a liquid the horizontal pressure is the same in all directions, and since the diameter of the cylinder may be indefinitely diminished, the vertical pressures must be equal for the same reason.

This may be illustrated experimentally by the apparatus shown in Fig. 32. Bend three glass tubes of about 5 mm diameter into the forms of tubes *I*, *II*, *III*. Fill the lower parts of the three tubes to the same height with mercury and immerse in a vessel of water, keeping the mouths of the three

tubes in a horizontal line. Now since the pressures at the open ends of the tubes are transmitted undiminished to the mercury surfaces, the difference in level of the mercury in the two arms of each tube is a measure of the pressure at the open end. *This difference of level is the same in each tube*, thus showing that in the horizontal plane through the mouths of the tubes the upward, downward and lateral forces due to these pressures are equal. On lowering the tubes to the bottom this difference in level increases, thus showing that the pressure increases as the depth increases.

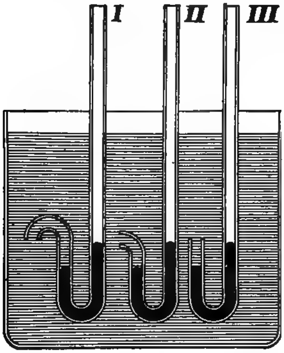


FIG. 32.

66. Free Surface of a Liquid at Rest, Horizontal. It may readily be shown that the free surface of a liquid at rest is horizontal. Let A (Fig. 33) be a particle of liquid of mass m, on a surface that is *not* horizontal. Then the force of gravity on the particle is mg in the direction AB. This can be resolved into a normal component AC, and a tangential component AD. The normal component simply produces pressure on the interior of the liquid and may be neglected. The tangential component AD is free to produce motion. Its value is

$$AD = mg \cos BAD \tag{161}$$

But by hypothesis the liquid is at rest; this cannot be true unless AD is zero. But since mg cannot be zero, $\cos BAD$ must be zero; that is, BAD is 90°, or the surface of a liquid at rest must be horizontal.

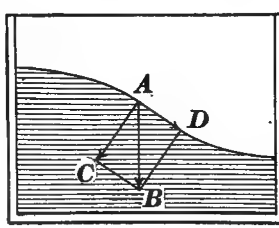


FIG. 33.

67. Pressure on an Immersed Surface due to the Weight of a Liquid. Consider a surface of area A immersed in a liquid of density d. We shall first find the force upon this surface due to the weight of the liquid. Let

the area be subdivided into a large number of very small areas, $a_1, a_2, a_3 \ldots\ldots a_n$; then

$$A = \Sigma a \tag{162}$$

Let us take one such elementary area a (Fig. 34) at a depth h from the surface, *and first consider this surface as horizontal.* The mass of liquid pressing upon this surface is the volume ah, multiplied by the density, or ahd, and the force f exerted upon this small surface is $ahdg$, *acting normally to the horizontal surface a.* But by (Arts. 63, 64) it has been shown that the forces due to fluid pressures *are always normal to the surface pressed upon,* and at a given point these forces are *the same in all directions.* Hence the force exerted upon the elementary area a is $ahdg$ regardless of the orientation of this area. The total force upon the surface A is therefore

$$\begin{aligned} F &= a_1h_1dg + a_2h_2dg + a_3h_3dg + \ldots\ldots + a_nh_ndg \\ &= (a_1h_1 + a_2h_2 + a_3h_3 + \ldots\ldots\ldots\ldots + a_nh_n)dg \end{aligned} \tag{163}$$

Let H be the distance from the center of area of the surface pressed upon, to the surface of the liquid. Then by equation (95) we have

$$H = \frac{\Sigma ah}{\Sigma a}$$

or

$$F = AHdg \tag{164}$$

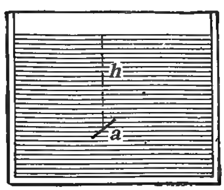


FIG. 34.

This is a general expression for the force due to the weight of a liquid upon any surface immersed in it, in any position whatever. The expression shows also that the force upon an immersed surface varies directly as the area of the surface, as the depth of the center of area below the surface of the liquid, as the density of the liquid, and as the acceleration due to gravity.

Equation (164) also gives an expression for the average pressure exerted by a liquid upon an immersed surface. For

$$P = \frac{F}{A} = Hdg \tag{165}$$

while the pressure p, at any point in the liquid, is obviously

$$p = \frac{f}{a} = hdg \tag{166}$$

68. Principle of Archimedes. *A solid immersed in a liquid is buoyed up by a force equal to the weight of the liquid displaced.* This principle is readily shown to be true by considering a vessel filled with liquid which is in equilibrium. This condition of affairs will not be disturbed if we suppose a certain part of the liquid to become solid without change of density. The forces acting on the solidified portion are its weight acting downward through its center of gravity, and the force exerted by the liquid upon its outer surface. Since this equilibrates the weight of the solidified portion, it follows that the resultant of all the forces due to the fluid must be directed upward through the center of gravity of the solidified portion, and must be equal to the weight of that portion. Consequently,

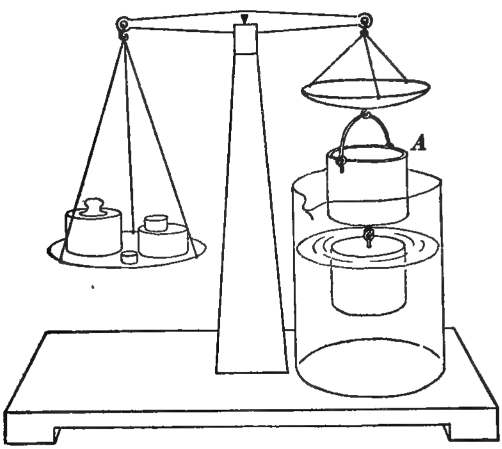


FIG. 35.

the buoyant force exerted upon any body immersed in a fluid (Fig. 35) is equal to the weight of the fluid displaced by the body, and acts through the center of gravity of the fluid displaced.

A result of this fact is that any body weighed in a fluid

weighs less than if weighed in air, by an amount equal to the weight of the fluid displaced. This seeming loss of weight is due to the buoyant force of the fluid in which the body is immersed. Hence to find the volume of any body of irregular outline it is only necessary to weigh it first in air and then in distilled water at 4° C. The apparent loss of weight is the weight of a volume of water, equal to the volume of the body, and if the weight be expressed in grams, the volume is at once obtained in cubic centimeters. In this way the number of cubic centimeters of iron in a handful of carpet tacks may be determined.

In accordance with the principle of Archimedes, it is necessary in accurate weighing to correct for the buoyancy of the air. This precaution is unnecessary in case the object weighed and the weights used have the same density; it is the more necessary, the smaller the density of the body weighed.

If we immerse a body in a fluid whose density is greater than that of the body, the body will displace more than its own weight of the fluid and will therefore rise until the weight of the fluid displaced just equals its own weight, when it floats. Hence *a floating body displaces its own weight of the liquid in which it floats.*

Finally, if we let m be the mass and w the weight of the body,

w_1 its apparent weight in the fluid,
m_1 its apparent mass when weighed in the fluid,
d the density of the body,
d_1 the density of the fluid,
v the volume of the body, and if we set the apparent loss of weight equal to the buoyant force of the fluid, we have

$$w - w_1 = mg - m_1 g = v d_1 g \tag{167}$$

69. Density and Specific Gravity. *The density of a body is its mass per unit volume.*[1] The specific gravity of a body is the ratio between the density of that body and the density of some substance assumed as a standard. For liquids and solids the standard is distilled water at a temperature of 4° C. In the

[1] *For experimental determination of density, see Manual, Exercises 28–30.*

C. G. S. system the numerical values for the density and the specific gravity of any substance are the same. In the English system, however, the density of water is 62.4 pounds per cubic foot, while its specific gravity is still unity. For gases hydrogen is usually assumed as the standard. If the mass and volume of a body be known, its density is given by the relation

$$d = \frac{m}{v}$$

Ordinarily the volume of a body is found in accordance with the principle of Archimedes, by finding the mass m of the body in air and its apparent mass m_1 in some fluid of known density d_1. The difference is the mass in grams of the fluid displaced by the body. The density of the body will then be given by the expression obtained in the previous article, on substituting for v its value m/d, or

$$d = \frac{m}{m - m_1} \cdot d_1 = \frac{w}{w - w_1} \cdot d_1 \tag{168}$$

In case the fluid is distilled water at 4° C, d_1 is taken as unity. If the weighing be done at some temperature other than 4° C, or the fluid be other than distilled water, then d_1 represents the density of the fluid used, *at the temperature at which the weighing was done.*

For bodies lighter than water, a sinker with a sharp point is hung from the left-hand pan, immersed in the standard fluid and counterpoised by small shot or fine sand. The body whose density is sought is placed in the left-hand pan and its mass m determined as usual. It is then fastened to the sinker by sticking it upon the sharp point immersed in the fluid, and equilibrium restored, either by adding masses to the left-hand pan, or removing them from the right-hand pan. The apparent change in mass represents the mass of the liquid displaced, or $m - m_1$. Whence, as before,

$$d = \frac{m}{m - m_1} \cdot d_1$$

For finding the density of liquids, the most accurate method is by means of the pyknometer, or "specific gravity bottle."

This is a small flask fitted with a ground glass stopper which is perforated throughout its length for the escape of superfluous fluid on filling the flask. The flask is first cleaned and dried and weighed empty. It is then filled with the liquid and weighed, and finally filled with the standard and weighed. These two weights, less the weight of the empty flask, give at once the masses of equal volumes of the substance and of the standard; whence we have

$$d = \frac{m}{v}$$

Other methods are frequently employed for determining the density of liquids, especially when great accuracy is not required.

70. Liquids in Communicating Tubes. Consider a bent tube (Fig. 36) containing two liquids which do not react chemically upon each other. When the system has come to rest, it will be found that the less dense liquid stands at a height h above a horizontal line through the junction ss' of the two liquids. The pressure exerted by this column is evidently balanced by the pressure due to the denser liquid whose height above the same surface is h_1. Let d and d_1 be the respective densities of the two liquids. The equation of equilibrium is then

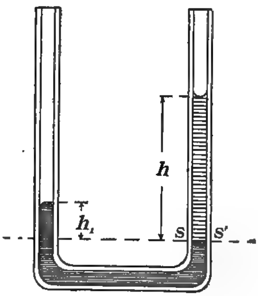


FIG. 36.

$$h_1 d_1 g = hdg \tag{169}$$

whence

$$\frac{d_1}{d} = \frac{h}{h_1} \tag{170}$$

or the heights of the two liquids above their common plane of separation vary inversely as their densities.

In case the liquids react chemically upon each other, the device shown in Fig. 37 may be used. The bent tube is inverted and the ends placed in cups containing the liquids of

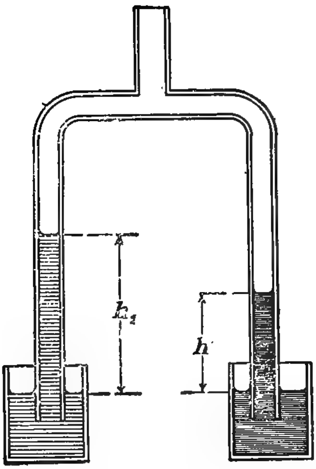


Fig. 37.

densities d and d_1. Through a short tube at the top the air may be partly removed from the tubes, producing a difference of pressure P between the air inside and outside the tube. This difference is balanced in each case by the rise of the liquid in the two tubes, and we have for equilibrium

$$h_1 d_1 g = hdg$$

as before.

If the density of one of the liquids be known, the density of the other may be computed at once.

Table IV

Densities of Various Bodies

Alcohol at 20° C	0.789	Iron, wrought	7.86
Aluminium	2.58	Lead	11.3
Brass (about)	8.5	Mercury at 0° C	13.595
Brick	2.1	Nickel	8.9
Copper	8.92	Oak	0.8
Cork	0.24	Paraffin	0.9
Diamond	3.52	Pine	0.5
Glass, common	2.6	Platinum	21.5
Glass, heavy flint	3.7	Quartz	2.65
Gold	19.3	Silver	10.53
Ice at 0° C	0.916	Tin	7.29
Iron, cast	7.4	Zinc	7.15

71. The Barometer. A barometer is an instrument for measuring the pressure of the atmosphere. It is made by taking a stout glass tube (Fig. 38) about 80 cm long and closed at one end, filling it with mercury, and inverting it with the open end under the surface of mercury in a shallow dish. When the system has come to rest, the mercury in the tube stands

about 76 cm above the level of the mercury in the dish, if the experiment be performed at sea level, in latitude 45°.

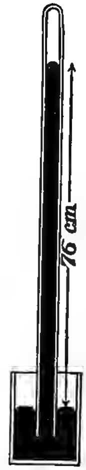


Fig. 38.

Now since pressure in fluids is transmitted undiminished in all directions, it follows that the weight per unit area of this column of mercury, 76 cm in height, must be balanced by the downward pressure of the air upon the surface of the mercury in the dish. If for any reason the atmospheric pressure change, the corresponding difference in the height of the mercury in the tube enables us to measure this change.

If the barometer be carried up the side of a mountain, or down into a mine, the elevation above, or the depression below, the sea level may be roughly determined from the difference in the barometric readings. At sea level a change of 11 m in level produces a change of 1 mm in the height of the barometer column. This rate of change is not constant, but diminishes as the elevation increases. Correction, however, should be made for differences in temperature at the various heights.

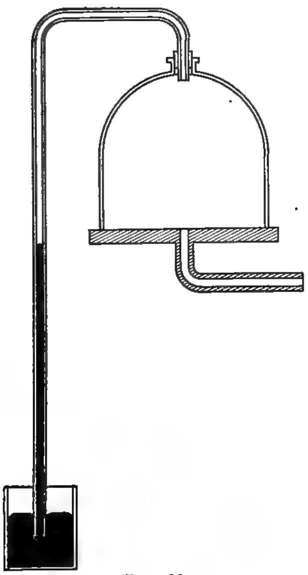


Fig. 39.

The pressure of the atmosphere as measured by the barometer is readily computed. Assume the tube (Fig. 38) to be of unit cross section and g to be $980\ \frac{\text{cm}}{\text{sec}^2}$. Then the volume of mercury supported is 76 cc; the mass is $76 \times 13.596 = 1033.296$ g, and the force is $1033.296 \times 980 = 1{,}012{,}630$ dynes. Hence the atmospheric pressure is 1,012,630 dynes per square centimeter. This pressure is called the *pressure of one atmosphere*.[1] In English units

[1] *For directions for the use of the barometer, see Manual, Exercise 15.*

the pressure of one atmosphere is approximately equal to a weight of 14.7 lb per square inch.

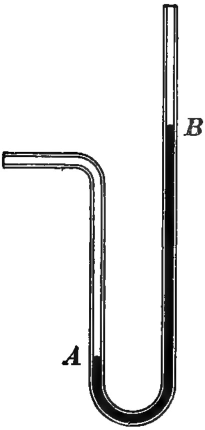


Fig. 40.

It should be noted that the barometric height is constantly changing, and that the standard height of 76 cm or 30 in is the average height for places at the *level of the sea, latitude 45° and at the temperature of 0° C.* Ann Arbor is 882 ft above sea level, and the mean barometric reading for the year 1901 was 29.03 in.

72. Manometers. Let a tube be attached to a bell jar and the jar be placed upon the plate of an air pump with the lower end of the tube dipping into a cup of mercury, as shown in Fig. 39. On withdrawing the air from the bell jar the mercury rises in the tube, thus furnishing a measure of the exhaustion as the pumping proceeds. Such an arrangement is called a *manometer.* For measuring slight differences in gaseous pressure, the second form shown in Fig. 40 is used, where the total pressure upon the gas in the horizontal arm is one atmosphere, plus the mercury column *AB*. A more sensitive form of this instrument is secured by substituting some light oil in place of mercury.

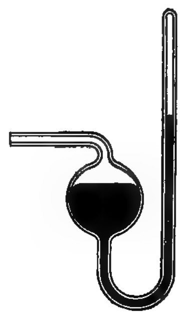


Fig. 41.

For the measurement of pressures amounting to several atmospheres the *compression manometer* (Fig. 41) is a convenient and compact form.

73. Pumps. In the action of the ordinary lifting pump (Fig. 42) we have an application of the pressure of the atmosphere. This pump consists of a piston *D*, working air tight in a cylinder *HK*, to which is attached a pipe of smaller diameter, closed by the valve *A*. Suppose the pump to be inserted in the water at *L*, and the piston to be at the bottom of the cylinder. On raising the piston the weight of the air closes the

valves B and C in the piston, and the air in the small pipe expands, raises the valve A, and passes in to fill the vacant space in the cylinder. The pressure of the air in the pipe is diminished and the hydrostatic pressure due to the weight of the atmosphere drives the water up into the small pipe. On the downward stroke the air in the cylinder closes the valve A, and escapes through the valves B and C into the outer air. By the second and succeeding strokes of the piston, the air below is still further rarefied, until the water rises above the valve A, passes above the piston through the valves B and C on the next downward stroke, and is finally lifted to S, where it flows out. Since the density of mercury is 13.6 times that of water, it follows that the pressure of the atmosphere will support a column of water 13.6 times the barometric height; or 30 in $\times 13.6 = 34$ ft. This would represent the maximum effect to be obtained from the pressure of the air; as a matter of fact an ordinary pump will not raise water much more than 26 ft.

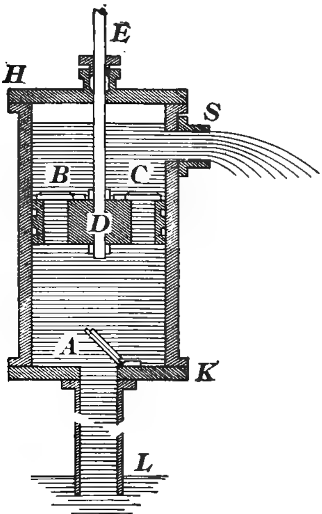


FIG. 42.

The *force pump* (Fig. 43) is provided with a solid piston L, and a second pipe is attached to the cylinder above the valve V_1. This pipe leads

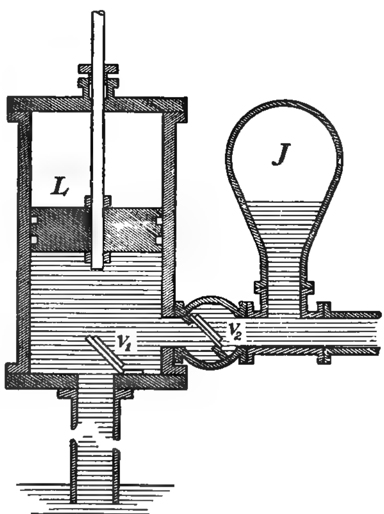


FIG. 43.

to an air chamber J, into which the water is driven on the down stroke of the piston, and retained there by a valve V_2. An outlet pipe fitted with a nozzle delivers the water driven out of the air chamber by the force of the compressed air; in other respects the working of the pump is similar to that of the lifting pump. The advantage of the air chamber is that it produces a steady stream from the nozzle instead of an intermittent one.

74. The Siphon. The siphon (Fig. 44) is an instrument for transferring liquids from one vessel to another at a lower level. It consists of a bent tube, with arms of unequal length filled with liquid and inverted with the shorter arm in the vessel from which the liquid is to be transferred. The distances a and b, from the surfaces of the liquid in the vessels to the highest point of the bend of the tube, represent the two columns of liquid in motion.

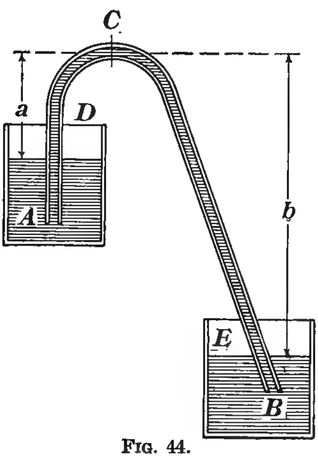


FIG. 44.

Let h be the height of the column of liquid which the atmospheric pressure will support, and d the density of the liquid. Then, while B is closed, the pressure at the level E inside the tube is

$$hdg + (b - a)\,dg \qquad (171)$$

while at the same level outside it is

$$hdg \qquad (172)$$

When B is opened, equilibrium is impossible and the liquid is forced out of the tube due to the difference in pressure

$$hdg + (b - a)\,dg - hdg = (b - a)\,dg \qquad (173)$$

If a equal b, the flow ceases. If a be greater than b, the liquid flows in the opposite direction.

75. The Air Pump. The action of the air pump is essentially the same as that of the common lifting pump.

In addition to the parts described (Art. 73) the air pump is fitted with an automatic valve *A* (Fig. 45), which opens at the up stroke of the piston and closes on each down stroke. This is effected by making the valve *A* in the form of a conical plug attached to a rod passing through the piston, which slides upon the rod with a small degree of friction. By this means the valve *A* is opened to admit air from the receiver, which by reason of its diminished density would be unable to lift the valve by its expansive force. A manometer *F* shows the degree of exhaustion in the receiver as the action of the pump proceeds. Owing to unavoidable wear and consequent leakage

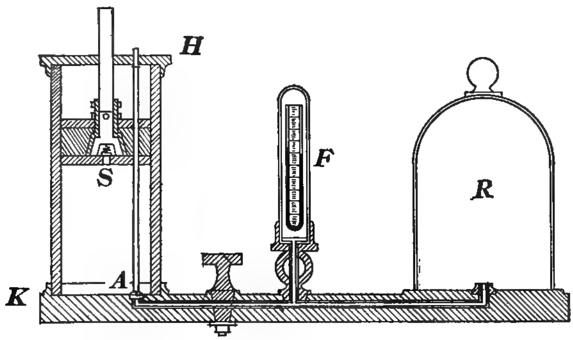


FIG. 45.

in the pump itself, it is impossible to secure a high degree of exhaustion with a mechanical air pump of the ordinary type.

For work requiring a high degree of exhaustion, as in the production of Geissler or Roentgen tubes, a mercury air pump is employed. Such pumps are usually made almost entirely of glass, and the piston is a column of mercury, which, on rising, fills a large bulb, expelling the air through a barometric manometer. On descending, the air from the vessel to be exhausted enters the bulb through a glass-mercury valve, and expanding, fills it at a much diminished pressure, only to be forced out as before by the liquid piston. By the use of such pumps pressures of one millionth of an atmosphere are readily attainable. For such purposes a variety of forms of the mercury pump have been devised, the idea in each case being to secure rapid automatic action.

76. Weight and Density of Air. Galileo satisfied himself that air had weight by weighing a glass globe first filled with air at ordinary pressure and then filled with air under high pressure. The experiment is ordinarily performed to-day by weighing a stout glass balloon when filled with air and then when the air has been exhausted. A liter of ordinary air, containing 0.04 % of carbon dioxide, under the standard conditions of 0° C and 76 cm pressure, weighs 1.293 g. Hence the density of air is 0.001293 g per cubic centimeter.

In the use of a delicate balance it is frequently necessary to correct the weighings for the buoyancy of the air in which the substance is weighed in accordance with the principle of Archimedes. Thus let w be the weight of the body weighed, w' the weight of the brass weights, and α and α' the buoyant force of the air upon the body and weights respectively. Then for equilibrium

$$w - \alpha = w' - \alpha'$$

or

$$w = w' + \alpha - \alpha' \qquad (174)$$

Thus the correction to be added to the weight of 25 g of cork, whose density is 0.25 g per cubic centimeter, is obtained as follows:

$$\alpha = \frac{25}{0.25} \times 0.001293 = 0.1293 \text{ g}$$

$$\alpha' = \frac{25}{8.5} \times 0.001293 = 0.0038 \text{ g}$$

Or the correction to be added is 125 mg.

77. Boyle's Law. Mention has already been made of the work necessary to compress a gas in a cylinder, and of the resistance offered by a gas to any force tending to *decrease* its volume. The relation between the volume of any mass of gas and the pressure exerted by the gas upon the walls of the containing vessel was first stated by *Robert Boyle* in 1662, although its far-reaching importance was more fully realized by *Mariotte*, a French physicist, who rediscovered the law, independently, in 1676.

Boyle's law states that, *at a constant temperature, the volume of a body of gas varies inversely as the pressure to which it is subjected.*[1] Thus if p_0 and v_0 be the initial pressure and volume of a mass of gas, and p and v the final pressure and volume, then

$$pv = p_0 v_0 = \textit{constant} \tag{175}$$

or *for a constant temperature, the product of the pressure and the volume is a constant.* Since the density of any body varies inversely as the volume, the law may be stated in another way, viz.:

$$\frac{p}{d} = \frac{p_0}{d_0} = \textit{constant} \tag{176}$$

or *the pressures are directly proportional to the densities.*

Careful experiments have shown that Boyle's law is only approximately true. It has been shown that all gases may be liquefied by the application of pressure and the reduction of the temperature. As the gas approaches the point of liquefaction, the variation from the law is most noticeable. Gases which can be liquefied by pressure at ordinary temperatures, such as chlorine, sulphur dioxide and carbon dioxide, can hardly be said to obey the law at all at these temperatures, while for gases like nitrogen, oxygen and hydrogen, the law is very nearly true at ordinary temperatures and for small differences in pressure.

An ideal or perfect gas would obey Boyle's law at all temperatures, and in general it may be said that the farther removed any gas is from its point of liquefaction, the more nearly it behaves as a perfect gas. This departure from the law by all gases when near the point of liquefaction has been attributed to the action of two causes: (*a*) *The attraction of the molecules of the gas for each other.* This would increase as the compression increases and therefore assist the compression. (*b*) *The size of the molecules,* which would tend to retard the compression of the gas. More complete formulae for the behavior of a gas under varying pressure and temperature have been proposed by van der Waals (Art. 219), Clausius and others.

[1] *For experimental verification of this law, see Manual, Exercise 16.*

CHAPTER VIII

FLUIDS IN MOTION

78. Velocity of Efflux. When a small opening is made in the side of a vessel filled with liquid at a distance h below the surface, the liquid flows out with a definite velocity v where

$$v^2 = 2\,gh$$

This formula for the velocity of a liquid issuing under pressure due to a head h is known as *Torricelli's theorem.* It may be demonstrated as follows: Suppose a particle of liquid of mass m to be situated in the surface of the liquid. Its potential energy with respect to the orifice is mgh. In passing from the surface of the liquid to the orifice it has fallen a distance h, and, if we neglect the viscosity of the liquid, its kinetic energy on emerging from the orifice must equal its potential energy at the beginning, or

$$\frac{mv^2}{2} = mgh \tag{177}$$

whence, as before,

$$v^2 = 2\,gh$$

If a be the area of the orifice, then V, the volume of liquid discharged in time t, would be

$$V = avt \tag{178}$$

In practice this rate of discharge is never reached. If the opening be a simple orifice in the side of the vessel, without mouthpiece of any sort, the volume of liquid discharged is about 0.62 of the theoretical value. This difference is due to the convergence of the lines of flow, producing a contraction of the jet just outside the orifice, whereby the actual cross section of the stream at this point is much reduced. By using a short

cylindrical mouthpiece, of length two or three times its diameter, and set flush with the inside of the vessel, the flow may be made 0.82 of the theoretical value; while by so shaping the mouthpiece as to conform most closely to the form of the contracted jet, a volume but little short of *avt* is attained.

79. Velocity of Effusion for Gases. Consider a vessel filled with gas, and having an orifice at one side of cross section a, from which the gas escapes with a velocity v. Let the density of the gas be d g/cm³, and the pressure be p dynes/cm², above that of the air. The volume of gas delivered in t sec will then be *avt* cc, and the mass *avtd* grams. The kinetic energy of the escaping gas will be $\frac{av^3td}{2}$ ergs. The gas is doing work upon the surrounding air by virtue of its expansion. If the outflow takes place so slowly that the gas is not cooled, this work is equal to the kinetic energy of the gas and is measured (Art. 35) by the product of the constant difference in pressure p and the increase of volume *avt* of the gas, or

$$\frac{av^3td}{2} = apvt \tag{179}$$

whence

$$v = \sqrt{\frac{2p}{d}} \tag{180}$$

or the velocity with which a gas effuses through an opening varies directly as the square root of the difference of pressure on the two sides of the opening and inversely as the square root of the density of the gas.

From this it follows that two gases effuse through the same opening, under the same difference of pressure, with velocities inversely as the square roots of their respective densities. Bunsen has utilized this principle to compare the densities of gases by observing the time required for the same volume of the various gases to effuse through the same opening under the same difference of pressure.

If the pressure p be expressed in terms of a column of the

gas extending h cm above the orifice, equation (180) then becomes

$$v=\sqrt{\frac{2\,ghd}{d}}=\sqrt{2\,gh} \tag{181}$$

*** 80. Flow of Liquids through Tubes.** If instead of a metal tube, an elastic tube of rubber be attached to the orifice of the discharging vessel, the efflux is the same as that by a rigid tube of the diameter assumed by the elastic tube, *so long as the flow is uniform.* If, however, the flow by any means be made to assume an intermittent or pulsating character, the discharge from the elastic tube is notably larger than from the rigid tube. The nature of the discharge is also modified, in that the stream from the rigid tube reproduces faithfully every feature of the pulsating impulse, while the elastic tube rapidly smooths out the inequalities of pressure, so that in an elastic tube of sufficient length the pulsation disappears entirely. This fact is of importance in explanation of the flow of the blood through the arteries, the coats of which are extremely elastic.

Again, the flow of liquids through tubes is much retarded on account of friction, not only among the particles of the liquid but between the liquid and the walls of the tube. This latter friction is much the more important of the two, and increases rapidly with the roughness of the walls of the tube. The flow of a river is greatest at the center of the stream, and at the surface of the water, where the effect of friction from the bed and banks is as small as possible.

In a vertical tube the liquid column breaks into a series of liquid masses fitting the tube more or less perfectly. These masses acquire an increasing velocity in their descent and act as liquid pistons fitting the interior of the tube. A partial vacuum results and the water is forced into the pipe more rapidly on this account. The effect of this exhaustion causes the noisy draught with which the last portions of water leave a wash basin or a bath tub, where the waste pipe is long and free. This action of vertical waste pipes is also liable to draw out the water from the siphon traps, and leave the way open for the entrance of

poisonous sewer gas. For this reason all waste pipes from basins, closets, bath tubs, etc., are now required to have separate vent to the outside air.

This reduction of pressure by liquids in vertical tubes is utilized in the Bunsen air pump, where a vertical column of water of more than 34 ft is made to exhaust the air from a receiver; the limit of the exhaustion being of course the pressure equal to the vapor tension of water at the existing temperature. In the Sprengel pump the liquid is mercury. This has two advantages: it requires a vertical column but 30 in long, and the vapor tension of mercury is practically negligible.

***81. Flow in Pipes of Variable Section.** In a tube of variable cross section the flow of liquids presents some interesting features. In Fig. 46 the variation of the pressure exerted by the fluid upon the walls of the tube is shown by the manometer tubes. It is thus seen that in a tube of variable cross section running full of liquid, the pressure is greatest in the widest part of the tube and least in the narrowest part. This somewhat surprising result is easily explained by considering the conditions of flow in the different parts of the tube. It is readily seen that with steady flow the velocity is greatest where the cross section of the stream is least, and *vice versa.* The liquid in passing from a wider to a narrower part of the tube must, therefore, undergo an acceleration, since for steady flow the same volume must pass any section in the pipe in the same time. To produce this acceleration the pressure on any section must be greater from *behind;* similarly, in passing from the narrower to the wider part of the tube, the velocity *decreases*, or the acceleration is *negative;* hence the pressure is greater from *before* than from *behind*, so that the pressure is greatest

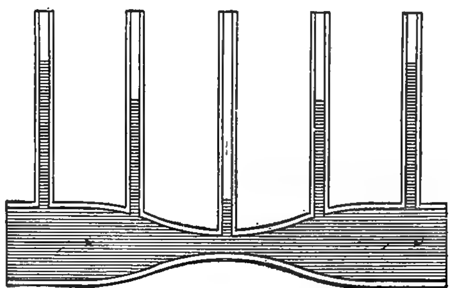


FIG. 46.

in the widest part of the tube and least in the narrowest part.

***82. Jet Pump.** The reduction of pressure within a contracted stream has been applied in the construction of many useful pieces of apparatus. In Fig. 47 is seen a common form of aspirator or jet pump as used in the laboratory. Water entering through the tube *E*, under hydrant pressure, passes through the constricted inner tube at *A*, and flows out at *D*. Owing to the small cross section of the stream at *A*, the velocity is very great, and the pressure is so much reduced that the air from the tube *B* is drawn along through the constricted portion in a torrent of small bubbles and carried down the tube *D*. With a well-constructed pump of this kind, a vacuum of about 5 cm of mercury may be obtained, with water from the city water mains. Obviously the pump will work equally well if the water enters at *B*, and the air through *E*.

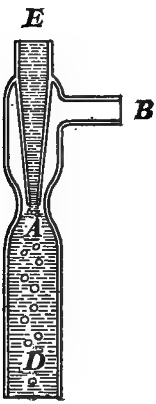


FIG. 47.

This pump has been adapted to numerous uses. The filter pump of the laboratory, the atomizer for spraying of perfumes or medicines, the injector in steam boilers, and the forced draught as used on locomotives are all different forms of this apparatus.

Problems

1. Find the weight on the bottom of a tank 10 ft square and 5 ft deep, when full of water. Find the force on one side of the same tank.

Ans. (*a*) 31,200 pounds.
(*b*) 7800 pounds.

2. A triangular plate is immersed vertically in water, with the vertex in the surface of the water and the base horizontal. The height and base of the triangle are each 50 cm. Find the force on the face of the plate.

Ans. 4083×10^4 dynes.

3. What is the density of a body whose mass is 678 g, if it weigh 235 g when immersed in a fluid whose density is 1.94 g per cubic centimeter?

Ans. 2.969 g per cm^3.

4. A piece of wood of density 0.6 g per cubic centimeter floats on water. The volume of the wood is 40 cc. What is the volume of the water displaced? *Ans.* 24 cm^3.

5. A body having a density of 2.35 g per cubic centimeter weighs 624 g when immersed in a liquid whose density is 0.827 g per cubic centimeter. What is the mass of the body? *Ans.* 962.84 g.

6. If the density of ice be 0.9179 g per cubic centimeter, and of sea water be 1.025 g per cubic centimeter, what portion of an iceberg is above water? *Ans.* 0.104.

7. A piece of silver and a piece of gold ore are suspended from the ends of a balance beam having equal arms. The balance is in equilibrium when the silver is immersed in alcohol (sp. gr. 0.85), and the gold in nitric acid (sp. gr. 1.5). If the specific gravities of the gold and silver be 19.3 and 10.5 respectively, find the ratio of their masses. *Ans.* As 1 to 1.004.

8. A U-tube is partly filled with water. How many centimeters of oil having a density of 0.79 g per cubic centimeter must be added in order to raise the water in one leg 4.5 cm above its first level? *Ans.* 11.392 cm.

9. Twenty-four cubic centimeters of gas at a pressure of 71 cm of mercury would have what volume under a pressure of 76 cm? *Ans.* 22.42 cm^3.

10. A liter of air under normal conditions of temperature and pressure weighs 1.293 g. Find the weight of the air in a liter flask when the barometer stands at 72 cm, the temperature being 0° C. *Ans.* 1.225 g.

11. To what depth must a diving bell 150 cm high be immersed in order that the water may rise 100 cm within it? Barometric reading 74 cm. *Ans.* 20.122 meters.

12. A glass tube used for sounding is 38.1 cm long and open at the lower end. The inside is covered with a soluble pigment, and the tube lowered to the bottom, in sea water, density 1.03 g per cubic centimeter. On raising it to the surface it is found that the water had entered the tube to a depth of 23.6 cm. Find the depth of the sea water. Barometric reading 74 cm. *Ans.* 16.3 meters.

13. A vessel filled with water has a circular orifice 6 cm in diameter, 298 cm vertically below the surface of the liquid. If the water be maintained at its initial depth, by supply from without, calculate the theoretical discharge per minute. *Ans.* 129.65×10^4 cm^3 per min.

14. A picture of mass 4 kilos is suspended in the ordinary way by a cord fastened to two hooks and passing over a smooth nail. The hooks are 45 cm apart and the cord is 120 cm long. Find the stretching force in the cord. *Ans.* 2.136 kilos.

MOLECULAR MECHANICS

CHAPTER IX

SURFACE PHENOMENA

83. Molecular Forces. By molecular forces are meant all those forces whose range of action is confined to insensible distances; that is, to distances comparable to the spaces between the individual molecules of a solid or a liquid. Under this head belong the forces of adhesion, cohesion, friction, viscosity, elasticity, capillarity, and surface tension; and although certain of these have been mentioned in previous topics, it seems proper to classify them here under the general term *molecular forces.*

The magnitude and importance of these forces are apt to be underestimated. It is owing to the action of molecular forces that any solid body is not only kept from falling to pieces of its own weight, but is able to resist the application of enormous stress as well. A clean glass tube cautiously lowered to the surface of clean water exhibits no attraction for the water, and causes no change in its upper surface so long as there exists any appreciable distance between them. But if the glass tube touch the surface, the water promptly runs up into the tube and clings to the outside, so that when the tube is withdrawn, a drop of water hangs to the lower end and the force of gravitation is unable to pull it off. Clearly we have here to do with forces, in comparison with which the force of gravitation is weak and insignificant.

The attraction between *unlike molecules*, as between those of water and glass, is called *adhesion;* that between *like molecules*, as between those of water and water, is called *cohesion.* These are in reality only different names for the same thing, viz., molecular attraction, and it is to be noted that this attraction is exhibited only so long as the substances are *in contact;* that is,

it acts through insensible or molecular distances. From a study of the behavior of gases we are led to believe that elasticity is due to molecular forces and molecular motions, and that the same is equally true of capillarity, surface tension and viscosity.

84. Adhesion and Cohesion. If two smooth, freshly cut surfaces of lead be firmly pressed together with a slight twisting motion, they cling together with considerable force, but having once been pulled apart they can be made to stick again only by application of sufficient force to bring the surfaces into close contact. A pair of glass plates, if highly polished, plane and free from dust, may be pressed together so firmly that it is impossible to separate them without rupture. That this is not due to the pressure of the air is shown by the fact that the plates cling together more firmly in a vacuum than in the open air, owing to the removal of the air film between the plates.

The adhesive action of glue, cement, mucilage and such substances renders it possible to unite two bodies so firmly that they break before separating. Dissimilar substances are united with difficulty, owing to their different rates of expansion or contraction when heated or cooled. Thus it is impossible to seal an iron or copper wire into a glass tube, since the metal and the glass have different rates of expansion; platinum, on the other hand, and certain alloys of nickel and iron may be sealed into glass, since they expand and contract at the same rate as the glass. For the same reason different kinds of glass frequently cannot be made to hold when sealed together.

Gases adhere to solids with great tenacity. It is almost impossible to free a glass tube from the adhering film of air, and consequently in the making of barometers, thermometers, and vacuum tubes of any description the air film is removed from the inner surface of the glass only by repeated heating and pumping.

The cohesive force of water is illustrated by hanging a clean, smooth glass plate to one arm of a balance so that it is horizontal, and bringing under it a jar of clean water. On touching the under side of the plate to the water, taking care to avoid air bubbles, it will be found necessary to add a relatively large weight to the opposite scale pan in order to pull the disk squarely

away from the water. When the disk is pulled away it is found that the *under surface is wet*, thus showing that the attraction between the solid and the liquid is greater than that between the particles of the liquid. This is true in all cases where *a liquid wets a solid.*

If mercury be used instead of water, it is found that *a greater force* will be needed to pull the plate away from the mercury and also that the *under side of the plate is now dry.* In the latter case we find that the attraction between the solid and the liquid is less than the attraction between the particles of the liquid, as is always the case *where a liquid does not wet a solid.*

85. Capillary Phenomena. If a solid be immersed in a liquid which wets it, the liquid rises about the sides of the solid, and the surface of the liquid is concave upward. If the solid be in the form of a tube, the liquid rises into the tube to a certain height, which varies inversely as the diameter of the tube, and forms as its upper surface in the tube a meniscus of liquid with its *concave side upward.* If the liquid do not wet the solid, it is depressed about the solid, or in case of a tube it is depressed to a certain depth, varying inversely as the diameter of the tube, below the level of the liquid in the vessel, and the surface of the meniscus is *convex upward.* Since phenomena of this class are most clearly shown in the case of fine, hairlike tubes, they are called *capillary phenomena*, from *capillus*, a hair. Fig. 48 shows the action when clean glass tubes are immersed in water and mercury respectively.

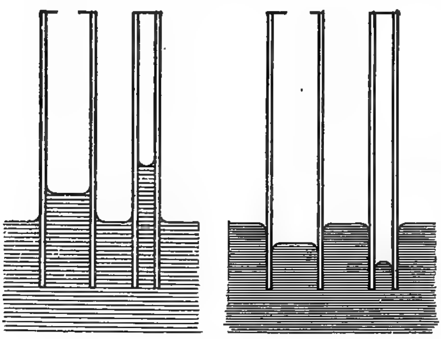


FIG. 48.

The principal facts concerning capillary phenomena are briefly these:

(*a*) In tubes of less than 2 mm in diameter the elevation or depression varies inversely as the diameter of the tube.

(*b*) The elevation or depression is independent of the pressure to which the liquid is subjected, and also independent of the thickness of the tube.

(*c*) An *increase* in the temperature of the liquid *causes a decrease* in the elevation or depression of the capillary column.

(*d*) The elevation or depression depends upon the nature of the liquid and tube in contact. Clean water on clean glass gives an elevation greater than that of any other liquid, while pure water in a steel tube or against a silver plate gives neither elevation nor depression.

Examples of capillary action are seen in the action of blotting paper absorbing ink, in the lamp wick supplying the flame with oil, and in the swelling of a tub or barrel if filled with water when about to fall to staves. A cotton or hemp rope, if wetted, absorbs water, increases in diameter and diminishes in length; at the same time the temperature of the rope rises through an interval of from 2° to 10° C. Workmen drive wedges of dry wood tightly into holes or slits cut in large stones and then wet the ends of the wedges. The increase in the size of the wedge is sufficient to burst the stone.

Besides the capillarity of liquids there seems to be an analogous phenomenon in the case of metals. Joseph Henry discovered that mercury would siphon through a bar of lead as water through a towel, and silver has been shown to pass into the pores of copper when the two metals are heated.

***86. Molecular Range.** In accordance with the assumption that molecular forces are exerted over insensible or molecular distances, it follows that each individual molecule becomes a center from which it exerts its molecular attractions and repulsions over the number of molecules included in its sphere of influence. Let the radius of this sphere be ϵ; then ϵ denotes the limit beyond which the molecule neither influences nor is influenced by its neighbors. Within this sphere, however, it is attracted equally on all sides and hence remains in equilibrium.

In order to determine the value of this quantity ϵ, Quincke employed a glass plate (Fig. 49), one half of which was coated

with a wedge-shaped layer, AC, of pure silver. On inserting the plate in water, with the silver film vertical (Fig. 50), the water rose against the glass above the level of the water in the dish to a definite height indicated by D.

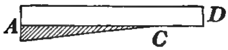


FIG. 49.

At the beginning of the silver film C, the elevation gradually fell away with increasing thickness of the silver, until at a point B, where the thickness reached a value of 0.000005 cm, the angle of contact became 90° and the capillary effect disappeared entirely. At this point the molecules of the glass ceased to influence the molecules of water through the silver, hence the value of ϵ, the molecular range, is commonly given as 0.000005 cm.

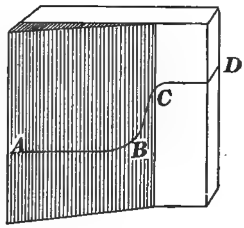


FIG. 50.

Recent investigations by Chamberlain show that this value is much too large, and that the true value of ϵ is 0.00000015 cm.

87. Surface Tension. It has been shown that within the limiting distance ϵ the molecules attract their neighbors and are attracted by them, and that a molecule situated in the body of a liquid will be in equilibrium by virtue of the equal attractions on all sides. Consider now a molecule nearer the surface of a liquid than the molecular range ϵ. In this case the horizontal attractions will be equal in all directions, but the vertical attractions are unequal; the resultant being an unbalanced component toward the interior of the liquid. At the surface of the liquid this resultant force, normal to the surface, reaches a maximum, and the mass of the liquid (Fig. 51) behaves as if surrounded by an elastic bag under stress, tending to contract indefinitely and compress the liquid into as small a volume as possible. This is due to what is known as *surface tension.*

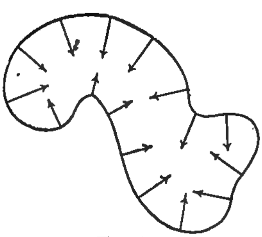


FIG. 51.

The surface of a liquid is therefore a seat of potential energy, since in order to force a molecule from the interior of the liquid

out into the surface film, work must be done in moving the molecule through the distance ϵ against the forces tending to draw it back into the interior. An increase in the area of the surface film, therefore, means an increase in its potential energy, and since potential energy always tends to become a minimum, it follows that if a liquid mass be freed from the action of other forces, it will assume a form such that its surface will be a minimum, and its volume a maximum; that is, it will assume the form of a sphere. This condition is readily realized by placing drops of olive oil in a mixture of alcohol and water of the same density as the oil. The drops are thus freed from the action of gravity and float as spherical globules in a medium of their own density. If by any device the globule be prevented from assuming a spherical form, it will still take a form presenting the minimum area consistent with the conditions imposed upon it.

Surface tension is exemplified in the shape of the dewdrop, in the forms of falling drops of liquid, and in the manufacture of shot, where molten lead, poured through a fine sieve at the top of a high tower, is broken up into small globules which harden as they fall through the air, and are caught in a tank of water beneath.

Again, small heavy bodies that are not wetted by a liquid may be placed upon the surface of the liquid and float upon the surface film. Thus needles may be made to float upon water, *so long as the film is not broken*, in which case they sink at once. The same principle is illustrated in the case of small insects which run upon the surface of the water, their slight weight being insufficient to break through the surface film.

88. Experiments on Surface Tension. If two small pieces of wood be floated near each other upon the surface of clean water and a drop of oil be touched to the water between them, they fly apart to the sides of the vessel, as though drawn by a spring. The addition of the oil reduces the surface tension of the liquid film at that point, and the water film tears apart.

If a plate of clean glass be wetted with clean water, the water will spread out into a thin layer over the entire plate. If the

plate be not absolutely clean, the water will gather up into irregular masses with rounded edges. If now a drop of alcohol be touched to the water layer, the water film will be seen to break at the point of contact, and gradually draw away from the alcohol drop, leaving a dry space on the plate around that point.

If a ring of stiff wire (Fig. 52) be dipped into a soap solution and withdrawn, a film of the solution will adhere to it for several minutes. In this film, which is really *two films* placed back to back, may be seen the motion of the liquid particles seeking new positions as the tension in the film changes. If we drop a loop of silk thread upon this film, it floats about freely upon it; if the film inside the loop be broken by touching it with a hot wire, the loop suddenly flies out into a circular form, showing that the tension in the film is equal in all directions. This circular loop still floats freely in the film, and behaves like an elastic hoop of steel.

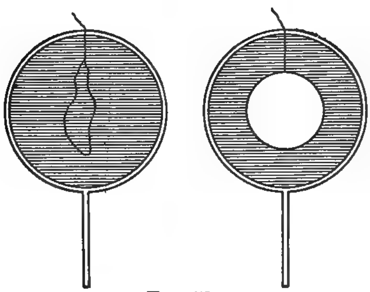


Fig. 52.

If small pieces of clean gum camphor be thrown upon the surface of clean water, the little particles begin a most lively and erratic motion. Each little piece spins with great vigor and at the same time sweeps over the surface, the larger ones gathering in the smaller ones. The gum camphor dissolves slowly in cold water, and the surface tension of the water film is thereby weakened. The spin is due to unequal dissolving on the surface of the camphor particle. If the water be warmed, it spins faster; if the surface be touched with a trace of oil, the motion ceases instantly.

If several small, clean, wood or paraffin balls be thrown upon clean water, they seem to attract each other and collect into a little cluster. If a number of similar balls be smoked with lampblack and then placed in the same dish, they also attract

each other, but the clean balls and the smoked balls seem to avoid each other.

A small cylinder of fine wire gauze, if immersed in water and lifted out in a horizontal position, retains the water in it and may be carried about the room. On breaking the water film at one point in the gauze by blowing upon it, the water begins to run out. The flow may be checked by shaking the water so as to restore the film, thus preventing the entrance of the air.

89. Measurement of Surface Tension. If a liquid exist in the form of a free film, then the two sides of the film exhibit surface tension in like degree and the film tends to contract indefinitely unless prevented by the application of an external force. If such a film be formed by dipping a rectangular frame of fine wire into a liquid and carefully raising it so as to keep its two sides vertical, the film tends to contract and draw the frame back into the liquid unless this contractile force be balanced by an external force. This affords a means of measuring surface tension.

If now frames of different width be taken and the forces needed to keep the frames in equilibrium be determined in each case, it will be found that the force is always *proportional to the width of the double film.* Let l represent the width of the frame, then the contractile force of the film is

$$F = T \times 2\,l \tag{182}$$

where T is a constant for any given liquid and is called the *surface tension* or the *capillary constant* of the liquid.[1] It is to be noted that in the case of surface tension, the term *tension* is used in the sense of *force per unit length* rather than *force per unit area* as usual. The unit of surface tension is one dyne per centimeter.

The following table shows the value of the surface tension in dynes per unit width of film, for the various substances mentioned. The values are mostly those given by Quincke.

[1] *For experimental determination of surface tension, see Manual, Exercises 31 and 32.*

TABLE V

SURFACE TENSIONS OF VARIOUS SUBSTANCES

Mercury against air	535.0
Water against air	81.0
Olive oil against air	36.9
Alcohol against air	25.5
Mercury against water	418.0
Olive oil against water	20.6
Turpentine against water	11.6

90. Capillary Action as Related to Surface Tension. Let a tube of radius r (Fig. 53) be inserted in a liquid of density d. Let the mean elevation of the liquid be h, and the angle of contact with the tube be α. Then the vertical component of the force due to surface tension T must be balanced by the weight of the liquid column of height h.

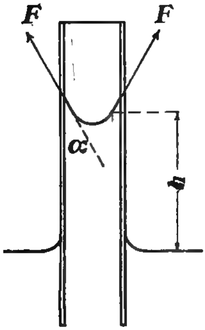


FIG. 53.

The width of the film around the tube is $2\pi r$ cm, and the total force in the direction indicated is $2\pi rT$; the vertical component is $2\pi rT\cos\alpha$.

The weight of the liquid supported is πr^2hdg; hence for equilibrium we have

$$\pi r^2hdg = 2\pi rT\cos\alpha \tag{183}$$

whence

$$h = \frac{2T}{rdg}\cos\alpha \tag{184}$$

For clean water on clean glass, the angle of contact is approximately zero, and

$$h = \frac{2T}{rdg} \tag{185}$$

For mercury, α is about 132°; h is negative, and the surface is depressed.

When we consider that the surface tension T *decreases* with

increase of temperature, it is seen that the above formula accounts for the inverse relation between capillary action and temperature.

For the elevation between two parallel plates distant u from each other the computation is similar to that above. The elevation or depression is

$$h = \frac{2\,T \cos \alpha}{udg} \tag{186}$$

or one half as great as for a tube of diameter u.

If two plates be joined at one edge and inserted in the water, the liquid rises high along the line of contact and falls off as the plates separate. The upper line of the fluid takes the form of an hyperbola, as shown in Fig. 54.

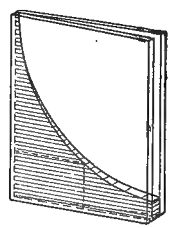


FIG. 54.

91. Angles of Contact. From the table in Art. 89, it is seen that there exists at the surface of separation between a liquid and a gas, or between two liquids, a surface stress or surface tension which is a constant for the same substances. Thus there exists in the surface film of olive oil in contact with air, a surface tension of 36.9 dynes per centimeter width of the film, while for water and air the surface tension is 81 dynes per centimeter width. If now a liquid be brought into contact with a second liquid in the presence of air, then for equilibrium, the three surface tensions should form a triangle of forces, and theoretically the angles between the forces should be constant. If, however, one of the forces should chance to be greater than the sum of the other two, then clearly no triangle is possible, and the system cannot come to equilibrium.

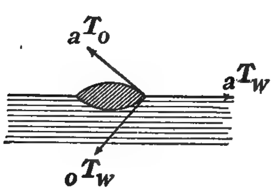


FIG. 55.

Thus in Fig. 55 is shown a drop of oil placed upon the surface of clean water. Then at any point upon the length l of the horizontal edge of the drop there are acting the three forces, $_aT_wl$ between the air and the water, $_aT_ol$ between the air

and oil, and ${}_oT_wl$ between the oil and water, directed in each case as indicated by the arrows. But from the table

$$
\begin{aligned}
{}_aT_w &= 81 \quad \text{dynes per centimeter width}\\
{}_aT_o &= 36.9 \text{ dynes per centimeter width}\\
{}_oT_w &= 20.6 \text{ dynes per centimeter width}
\end{aligned}
$$

hence

$$ {}_aT_w > {}_aT_o + {}_oT_w \tag{187} $$

This shows that when a drop of oil is placed upon clean water, the surface tension between air and water is so great as to overbalance the other two surface tensions combined, and the oil is dragged out in all directions, forming a film of infinitesimal thickness over the entire surface of the water.

If a liquid meet a solid in the presence of air, it will in general meet it in a definite angle which is constant for the two substances. This angle is called the *angle of contact*, and depends upon the nature of the substances in question. For pure water on clean glass the angle of contact is approximately zero. For pure water on clean steel or clean silver the angle of contact is about 90°. For clean mercury against clean glass it is about 132°.

*** 92. Behavior of Films.** If a film of soap solution be made to assume a curved form, there will always result a normal pressure directed towards the concave side. It may be shown mathematically that for a single cylindrical film, of radius R, and surface tension T, the normal pressure toward the curved side is given by the equation

$$ P = \frac{T}{R} \tag{188} $$

In words this equation says that the normal pressure in a curved film is directly proportional to the surface tension T and inversely as the radius of the film R. In general the curvature of any surface at any point may be expressed in terms of two radii of curvature, the planes of curvature being at right angles to each other. If R_1 and R_2 be these radii, then

the normal pressure of any curved film is the sum of that due to each curvature separately, or

$$P = T\left(\frac{1}{R_1} + \frac{1}{R_2}\right) \tag{189}$$

In a soap bubble the two radii are equal, and there are also two films back to back, hence the normal pressure exerted upon the air enclosed in a bubble is

$$P = 4\frac{T}{R} \tag{190}$$

If the film is free to the air on both sides, the normal pressure must be zero. In a curved film this is possible only if

$$\left(\frac{1}{R_1} + \frac{1}{R_2}\right) = 0 \tag{191}$$

This means that

$$R_1 = - R_2 \tag{192}$$

or that the radii are equal and on opposite sides of the film; that is, the film is saddle-shaped.

Again, since this normal pressure is directed toward the concave side and varies inversely as the radius, we are able to understand the motion of drops of water and of mercury in conical tubes. The water will move toward the smaller end of the tube; that is, in the direction of the greater normal pressure. For the same reason, the mercury globule will move toward the larger end of the tube.

If stout wire frames, representing the outlines of geometrical figures, be dipped into soap solution, a large number of curious and beautiful film figures result. Whenever three such film surfaces meet along a line, the included angles will each be 120°, since the three forces are all equal.

CHAPTER X

SOLUTION AND DIFFUSION

93. Solution. Closely allied to the phenomena of capillarity and surface tension are the phenomena of solution. Ostwald defines solutions as "homogeneous mixtures which cannot be separated into their constituent parts by mechanical means." Gases have unlimited power of solution. One gas dissolves in another in all proportions so long as they do not unite chemically, and the homogeneous mixture manifests the sum of the properties of the two constituents.

Liquids dissolve gases without exception, although the readiness with which such solution occurs varies greatly with the nature and condition of the substances. In a true solution of a gas in a liquid, the gas may be entirely removed from the liquid by diminishing the pressure or by raising the temperature, and in such solutions the quantity of gas dissolved by a given mass of liquid is proportional to the pressure to which the gas is subjected. Examples of this sort of solution are found in solutions of carbon dioxide, of air, or of ammonia gas in water. In other cases, as in the solution of hydrochloric acid gas in water, the dissolved gas is not entirely removed from the liquid, and it is assumed that a chemical change has resulted.

The solution of one liquid in another occurs in many cases, but is dependent upon the nature of the substances. Here also there are two distinct classes of solution. Thus some liquids, as alcohol and water, dissolve in all proportions, forming a homogeneous mixture. Ether and water, on the other hand, dissolve in each other, *but in limited proportions.* Thus water will dissolve about ten per cent of ether, but if more ether be added, the excess remains undissolved. Ether will dissolve about three

per cent of water, but beyond this the water remains separate from the ether.

A third division contains those liquids that do not dissolve in each other at all. This is a relatively small number, since even those liquids that seem insoluble yet leave traces in the solvent. Thus the fact that water when shaken with the volatile oils retains the characteristic odor of those oils seems to show that even here solution has occurred in slight degree.

Again, mixtures of insoluble liquids, as water and the fixed oils, may be made by the addition of some such substance as gum acacia or gum tragacanth, in which the oils are broken up into exceedingly small globules that float in the water. Such mixtures are called *emulsions.* Milk is a natural emulsion. Some emulsions separate on standing or when subjected to mechanical action, as seen in the separation of cream from milk.

94. Solution of Solids. Many solids when immersed in a liquid gradually disappear and form a new homogeneous liquid. The solid is said to dissolve in the liquid, and the new liquid is called the *solution.* A liquid that dissolves a solid is called the *solvent.* Many salts are soluble in water. The quantity of a substance in solution may vary from zero up to a certain limit, beyond which the solution has no further action upon the substance in question. Such a solution is said to be *saturated.* The amount of a solid that may be dissolved in any solvent varies with the temperature. If the temperature of a saturated solution in contact with its salt be changed, either some of the dissolved solid separates out or more of the undissolved solid goes into solution. Generally a solvent will dissolve more of a solid when hot than when cold although there are exceptions to the rule. Thus the solubility of sodium sulphate *increases* with the temperature up to 33° C, but beyond that temperature its solubility *decreases.*

If a solution, either by evaporation or reduction of temperature, be made to contain more than its normal quantity of a solid, it is said to be *supersaturated.* If a particle of the undissolved solid be dropped into the supersaturated solution, the excess of solid in the solution crystallizes out at once (Art. 195).

95. Free Diffusion of Gases. Dalton's Law. Let a tall glass jar be inverted and filled by upward displacement with illuminating gas, and placed upon a similar jar filled with air, with the mouths of the jars together. We shall thus have the two jars filled with separate gases, and since the lighter gas is on top, no mingling of the gases can be produced by the action of gravity. If now the jars be left in position for a quarter of an hour, we shall find, on testing the contents with a lighted splinter, that there is an explosive mixture of illuminating gas and air in each jar. This illustrates diffusion of gases, and this result can be explained only on the hypothesis that the molecules of the gases are in motion and that they have therefore wandered through the entire space, the heavier gas rising into the upper jar and the lighter gas descending into the lower jar.

After the gases are uniformly diffused it will be found that the pressure exerted by the mixture is the sum of the pressures exerted by its constituent parts. Thus if we allow 10 volumes of illuminating gas and 15 volumes of air each at atmospheric pressure p to diffuse uniformly, without change of temperature, through a space of 25 volumes, then, according to Boyle's law, the air would exert a pressures of $15p/25$, and the gas $10p/25$, and their combined pressure would equal that of the atmosphere outside. This important relation was first established by Dalton and is known as *Dalton's law*. It may be stated as follows: *A mixture of two or more gases having no chemical action upon each other exerts a pressure equal to the sum of the pressures which would be exerted by each of the constituent gases separately if allowed to fill the containing vessel alone at the given temperature.* It thus appears that each gas behaves as if no other gas were present, the only effect being a slightly diminished rate of diffusion, owing to the mutual molecular collisions. In general, the properties of such a mixture are found to be the sum of the properties of the various gases composing the mixture.

96. Diffusion of Gases through Porous Partitions. Atmolysis. A tube, Fig. 56, is partly filled with water, and the right arm closed with a porous cup. Over the porous cup is lowered an inverted beaker filled with hydrogen or illuminating gas.

The water sinks in the right arm and rises in the left, showing an increase of pressure in the cup. After the beaker has remained in position for a minute or two, suddenly remove it. The water now *rises* in the right-hand tube and is *depressed* in the left. The explanation of these two experiments is very simple. In the first, the lighter illuminating gas diffuses *inward* more rapidly than the air diffuses *outward*, and an increase of pressure in the cup results. In the second case the gas now inside the cup, being lighter than the air outside, diffuses *outward* more rapidly than the air diffuses *inward*, causing a reduction of pressure in the cup.

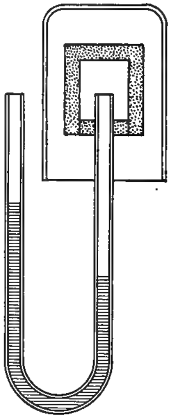


FIG. 56.

The differences in rates of diffusion for different gases have been utilized for separating a gaseous mixture into its constituent parts. Thus if we pass a mixture of hydrogen, nitrogen and oxygen through a porous tube made from the stems of clay tobacco pipes, and maintain a vacuum about the outside of the tube, we shall find that the hydrogen, being the lightest, will diffuse most rapidly through the walls of the tube, leaving the nitrogen and oxygen behind. Of course some nitrogen and some oxygen escape also, but the mixture transmitted by the tube is relatively richer in the heavier constituents, as indicated by equation (180). Rayleigh and Ramsay were able by this means to separate argon from atmospheric nitrogen. This process of separation of gases was first used by Graham, and was called by him *atmolysis*.

97. Diffusion of Gases through India Rubber, and through Red-hot Metals. In 1831 Mitchell observed that toy balloons made of india rubber collapsed much sooner when inflated with carbonic acid than when filled with air or even with hydrogen. Graham, who studied the phenomenon, found that while one volume of nitrogen would pass through a sheet of india rubber in a given time, 2.56 volumes of oxygen, 5.5 volumes of hydrogen and 13.58 volumes of carbonic acid would pass through in

the same time. Also the speed with which gases pass through rubber increases very rapidly with increase of temperature.

These remarkable facts do not seem to be connected in any intimate way with the transmission of gases through porous partitions. No simple relation connects the densities of the gases with their speeds of diffusion, as in the case of porous septa, and the process of transmission appears to be an entirely different one. The most probable explanation seems to be that the rubber is capable of absorbing and retaining certain amounts of the various gases with which it comes in contact, the amounts increasing rapidly as the pressure increases. After the surface layer of the rubber has become saturated, as it were, with the gas in question, this condition is then passed on from layer to layer of the rubber, until the outside layer is reached. Here, since the pressure is less, the rubber is unable to retain all the absorbed gas, and some of it escapes into the adjacent space.

Red-hot metals transmit gases with great facility. The poisonous carbon monoxide passes freely through cast iron at red heat, and from the red-hot cast iron coal stove this deadly gas leaks into the room almost as water from a sieve. Hydrogen diffuses through a red-hot platinum tube, whose walls are 1.1 mm thick, at the rate of *490 cc per minute for every square centimeter of surface.* Silver at high temperatures transmits oxygen readily. A glowing tube of palladium, through which is carried a mixture of CO and H, separates these gases completely, the hydrogen being transmitted and the carbon monoxide being retained. It thus appears that glowing platinum and palladium act as "semi-permeable membranes" for certain gases, just as certain solid substances do for liquids, in that they allow some substances to pass freely and refuse transmission to others. This peculiarity is of great importance in osmotic phenomena, as we shall see later. In all cases of such transmission through rubber or glowing metals, we seem to have to do with a species of solution of the gas on the one side of the partition, and of evaporation of the gas from the other side. The same process seems to account for similar behavior in the case of gases and liquid films.

98. Free Diffusion of Liquids. If two liquids that do not react chemically upon each other be left in contact with each other, they will of themselves begin to mingle at once, and continue until they form one homogeneous liquid throughout. Thus, if a solution of copper sulphate be placed in the bottom of a small jar, and carefully covered with distilled water, so the line of separation is well defined, and the jar be left undisturbed for a few days, we shall see that the blue color of the copper sulphate has risen into the clear water above, and that the line of demarcation is no longer sharp between the liquids. The color of the copper salt at the bottom has also become slightly less dense than at first, and the two liquids seem tending toward a uniform color.

This process is called *diffusion*, and while resembling the related phenomenon in gases, its progress in liquids is exceedingly slow. For example, if the jar containing the copper sulphate in the above example be made a meter high, the lower half filled with the solution and upper half containing pure water, it would take more than ten years for the solution to assume a uniform color throughout. If the jar were one centimeter high, it would require about ten hours, the time for uniform diffusion varying as the square of the length of the liquid column.

The speed at which a given solution will diffuse through the pure solvent depends upon the nature of the salt and of the solvent, upon the temperature, and to a slight degree upon the strength of the solution. From extended experiments it has been found that those salts having the highest electrical conductivity have also the highest velocity of diffusion.

99. Diffusion through Membranes. Osmosis, Crystalloids and Colloids. As we have seen, if two solutions of different strength be brought into contact, a condition of equilibrium cannot, in general, be maintained. A movement of the dissolved substance sets in *from the concentrated to the dilute* solution, and continues until it is uniformly distributed throughout the liquid. If, however, we enclose the solution in a vessel fitted with a manometer tube, and provided with a bottom of some

porous substance, as parchment or animal membrane, and immerse the whole in pure water, the process is very different.

The porous membrane does not allow as easy transmission to the molecules of the dissolved substance as to the molecules of the solvent. As a result, the solvent crowds in through the membrane and creates an internal pressure, as shown by the rise of liquid in the manometer. This crowding in of the solvent continues until the pressure reaches a definite value, depending upon the strength of the solution. After this, the tendency of the molecules of the solvent to enter the cell seems to be balanced by the internal pressure, and equilibrium ensues.

This unequal diffusion through porous septa is called *osmosis*, and the membrane is termed a *semi-permeable membrane*, if it completely prevents the passage of the dissolved substance. The limiting pressure beyond which no more of the solvent enters the cell, is called the *osmotic pressure* for the substance, *at that temperature* and *for that concentration.*

The phenomenon may be illustrated by the following experiment. A conical vessel attached to a long tube is closed at its larger end by a piece of bladder or parchment firmly tied on. When the vessel is filled with sugar solution to the lower end of the tube, and immersed in a vessel containing water, as shown in Fig. 57, the liquid in the tube rises to a considerable height above the level of the water in the outer vessel.

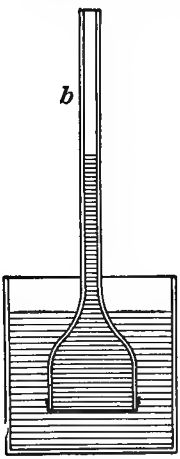


FIG. 57.

Through the experiments of Pfeffer it was discovered that the best results are to be obtained by attaching the tube to a closed clay cell, the pores of which are filled with a precipitate of copper ferrocyanide. The precipitate is pervious to water but impervious to the dissolved substance. With such a cell, filled with a 3.3 per cent solution of potassium nitrate, Pfeffer obtained an osmotic pressure of 436.8 cm of mercury, *or more than 5.7 atmospheres.*

Those substances which pass through animal membranes

most readily, such as mineral acids and neutral salts, are generally known in the crystalline form and hence have been called *crystalloids*. Substances like gums, tannin, albumen, starch, etc., which are amorphous in character and do not pass through so readily are called *colloids*. The effects of crystalloids when dissolved in water are very marked in the changes produced in the properties of the solvent. Thus any crystalloid dissolved in water *diminishes its vapor pressure, lowers its freezing point and raises its boiling point*. Colloids, on the other hand, when dissolved in water produce scarcely any such effects. Colloidal solutions in general represent loose mechanical mixtures from which in many cases the substance held in solution may be precipitated by a slight trace of acid or alkali. When mixed with small quantities of water the colloids form jellies, in some of which the structure is so coarse as to be visible under the microscope. This is notably so of the colloidal solutions of the salts of gold, in which the suspended particles of gold form the objects for ultra-microscopic vision by means of transverse illumination. While many of these colloidal jellies transmit crystalloids almost as readily as pure water, they offer great resistance to the diffusion of other colloids.

100. Osmotic Pressure. We have seen that the entrance of water through the semi-permeable membrane into the osmotic cell may be prevented by subjecting the enclosed solution to sufficient pressure. This pressure is called the *osmotic pressure*, for the substance in question under the conditions of the experiment, and is of great importance since the properties of the solution, such as its vapor pressure, its boiling and freezing points, are immediately calculable as soon as this pressure has been determined. Osmotic pressure is also intimately related to the transmission of fluids in the living cell of plant or animal tissues. Thus it has been shown that such a cell when placed in concentrated salt solutions, has its liquid contents diminished by the removal of water; if placed in a solution whose osmotic pressure is less than that of the cell, the cell and its contents are distended by the addition of water.

Oil globules of extreme smallness floating in water tend

more readily to pass bodily through the pores of animal membranes, if some alkali be mixed with the water, since they have a soapy covering and are thus passed through the pores like so much water. (Daniell.)

The following conclusions concerning osmotic pressure are stated by Ostwald as reasonably well determined for dilute solutions:

(*a*) The osmotic pressure depends upon the nature of the substance.

(*b*) The osmotic pressure is proportional to the concentration of the solution, or inversely proportional to the volume in which a definite mass of the dissolved substance is contained.

(*c*) The pressure for a given concentration is proportional to the absolute temperature.

(*d*) Quantities of dissolved substances not electrolytes which are in the ratio of their molecular weights exert equal pressures at equal temperatures.

(*e*) The pressure is independent of the nature of the membrane provided the membrane be impervious to the dissolved substance.

***101. Dialysis.** The division of substances into *crystalloids* and *colloids* has already been mentioned. A characteristic property of colloids is, that while they pass through porous septa with difficulty *themselves*, they offer no marked resistance to the diffusion of *crystalloids*, but are more or less impervious to other colloids. A result of this property is, that if a mixture of crystalloid and colloid substances be separated from pure water, by a colloidal membrane of a different kind, the crystalloids will soon diffuse into the water, while the colloids will remain behind.

This principle has been applied to the separation of crystalloidal poisons from a heterogeneous mass of organic matter in which their presence is suspected. The apparatus as employed by Graham consists of a hoop, over one side of which has been stretched a piece of bladder or parchment paper, put on wet and held in place by a string. This is slipped inside a

second hoop, thus forming a shallow dish with a colloidal bottom. This is now floated in a vessel containing distilled water and the substances to be examined are placed in a thin layer on the membranous bottom. If crystalloids be present, they will diffuse through into the distilled water, while the colloids are left behind. The water in the lower vessel may then be examined for the suspected poisons by ordinary analysis. The floating dish is called a *dialyzer*, and the process is termed *dialysis*.

Problems

1. How high will water stand above hydrostatic level in a tube 0.057 cm in diameter, assuming the angle of contact to be very small? *Ans.* 5.8 cm.

2. How much will the level of mercury be depressed in a glass tube 0.067 cm in diameter, the surface tension being taken as 540 dynes per centimeter and the angle of contact as 135°? *Ans.* 1.711 cm.

3. What is the pressure within a soap bubble 12.5 cm in diameter if the surface tension of the liquid be taken as 80 dynes per centimeter, and what is the total force exerted by the film on the gas within?

Ans. (*a*) 51.2 dynes per cm^2.
(*b*) 25,133 dynes.

SOUND

ORIGIN AND PROPAGATION

CHAPTER XI

NATURE OF SOUND

102. Definitions. Under sound are studied those manifestations of energy which primarily appeal to the ear. This branch of physics has to do with the study of vibratory motion in ponderable, elastic media. The phenomena of sound differ from those of heat, light and electricity in this respect, that the cause of the phenomena is definitely known, while in heat, light and electricity, the cause is assumed.

In common language the word *sound* is used in two distinct senses. It may mean the sensations reported to the brain by the auditory nerves, or it may refer to the external cause of those sensations. The psychologist and physiologist are concerned with the phenomena of sense perception. The physicist is interested in the external disturbance *producing the sense perception;* the conditions of its origin, its mode of propagation, and the variations in the nature of the disturbance corresponding to certain differences in the sensation produced. Accordingly sound is defined by the physicist as *that form of vibratory motion which may be perceived by the ear.* According to this definition sound may exist entirely independent of an ear to perceive or a brain to comprehend.

103. Origin of Sound. Sound originates in a vibrating body. "Sound and movement," says Blaserna, "are so correlated that one is strong when the other is strong, one diminishes as the other diminishes, and the one stops when the other stops." A

guitar string plucked aside and released gives a musical note, and at the same time seems to spread out into a broad band with a hazy outline, which diminishes to the original size of the string as the sound dies away. The tremulous motion of a bell may be perceived by placing the hand upon it while it is sounding. A long glass tube, if grasped by the middle and rubbed with a moistened cloth, gives forth a soft musical note, while the vibratory motion may be plainly felt by the hand. The air column in an ordinary tin whistle is thrown into vibratory motion when the whistle is blown. The whistle may even be blown by water and the instrument will give forth a soft clear note, produced by the vibration of the stream of water.

104. Wave Motion. We have seen (Art. 33) that a simple harmonic motion compounded with a uniform motion in a straight line produces a "sine curve." Such a curve (Fig. 17 *bis*),

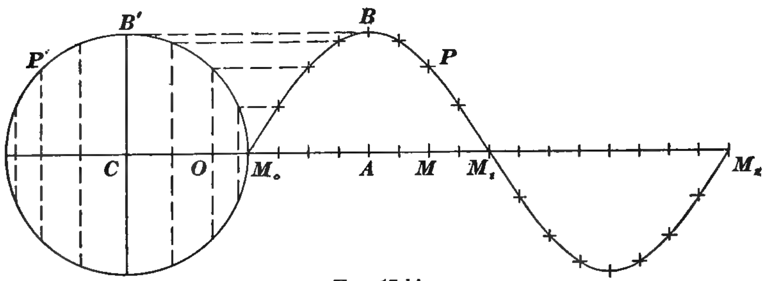


FIG. 17 *bis*.

may be regarded as arising from a series of equidistant particles arranged along the axis M_0M_2, each executing simple harmonic motion at right angles to this axis, but having a common difference of phase; that is, each particle to the right of M_0 crosses the axis going in the positive direction, $\frac{1}{16}$ *of a period later than its left-hand neighbor.* The result of such a system of moving particles is the sinuous form, known as a *sine curve* or *simple wave.*

During the time T, required for any particle to describe a complete vibration, the disturbance will have run to the right, the distance from M_0 to M_2; so that when M_0 begins its *second*

vibration, M_2 begins its *first.* Since M_2 marks the extreme advance of the wave disturbance at this instant, it may be said to lie on the *wave front.* By this definition special attention is directed to the limiting position reached by the disturbance. However, we frequently use the term *wave front* when speaking of an extended series of moving particles in space and, regardless of the limit reached by the disturbance, we are accustomed to say that a wave front is the *continuous locus of all points in the same phase of vibration.*

The distance M_0M_2 is called a *wave length* λ, and is related to the *velocity*, V, of *the wave*, and the time of vibration, T, of the particle, by the equation

$$\lambda = VT = \frac{V}{n} \tag{193}$$

where n denotes the *frequency* or the *number of vibrations* made by a particle *in one second.*

In the case considered the particles are conceived to vibrate in paths at right angles to the line of propagation of the wave; such a wave is termed a *transverse wave.* Examples are seen in the waves produced in a soft rope when one end is vibrated quickly and regularly in a line at right angles to its length.

If the particles vibrate back and forth *in* the line of propagation of the wave, the wave is called a *longitudinal wave.* Such a wave is produced by suspending a long spiral coil of brass wire between two fixed supports and producing a compression in it by forcing the spirals at one point closer together. The compression runs the entire length of the coil.

105. Characteristics of Wave Motion. The fundamental characteristic in wave motion is the continuous handing on, from point to point, in an elastic medium, of a periodic disturbance maintained at the source. Such a disturbance produces a series of waves which follow each other at definite intervals and which constitute what is known as a wave train. It is to be observed that wave motion *transfers energy* from one point to another, by means of the *motion of the particles of the medium.*

The individual particles oscillate about their positions of

rest, while the *wave form* runs forward. The distinction between the motion of the particle and the motion of the wave is of fundamental importance. It may perhaps be best illustrated by the motion of a wave passing over a field of grain. The wave form runs forward while the individual head of grain simply swings back and forth in a plane parallel with the line of propagation.

Other characteristics of wave motion may be noted as follows:

The disturbance requires time to travel from one point to another.

A medium is required for transmission of the disturbance.

Waves are *reflected* on meeting an obstacle, the angles of incidence and reflection being equal.

The direction of the wave is changed, that is, the wave is *refracted*, on entering a medium in which the speed of propagation is different.

Two systems of waves may be added together so as to reënforce each other, if crest meet crest and trough meet trough; or they may be added so as to annul one another and produce rest, when the crests of one system meet the troughs of the other. This is known as the *interference* of wave systems.

Waves, on meeting small obstacles, bend round corners and behind the obstacles, form a series of points of maximum and minimum disturbance. This phenomenon is known as *diffraction.*

Two waves may differ from each other in *amplitude*, in *wave length* or in *wave form.*

106. Characteristics of Sound. A sounding body transmits its vibrations to the surrounding medium and a system of sound waves is the result. The waves of sound are longitudinal in character, the vibrations of the particle taking place in the line of propagation. A comparison of the characteristics of wave motion with those of sound reveals the following points of similarity.

Sound requires time for its transmission from point to point; in other words, sound travels with a definite velocity, which depends upon the nature and condition of the medium. Exam-

ples of this finite velocity of sound are seen in the interval elapsing between the flash of a distant gun and the report of the discharge; or between the appearance of the puff of steam and the sound of the whistle of a distant locomotive, or in the case of the distant woodman who has his ax raised for a second blow before the sound of the former one reaches us.

Sound requires a continuous, elastic and ponderable medium. Sound is not transmitted in a vacuum. An electric bell suspended by rubber cords from the inside of a bell jar is silenced on the exhaustion of the air from the jar, although the clapper may still be seen to be moving. On readmitting the air the sound is heard again. Unless the medium be elastic the energy originally imparted to the medium by the vibrating body is largely lost in heat, as in the case of impact between inelastic solids. The medium must also be ponderable; that is, it must have weight. Sound is a *material* phenomenon. It is produced by the vibrations of material particles, and requires a material medium for its transmission. Liquids transmit sound better than gases, and solids better than either.

Sound is reflected. Whispering galleries are simply large rooms in which the curved surfaces of the walls or roof serve as mirrors for the concentration of the sound waves at particular points. Speaking tubes are devices for preventing the dissipation of sonorous energy, by confining it to a certain narrow space, through which it passes by repeated reflections from the sides of the tube. Echoes are examples of the reflection of sound by the walls of the building, by hills or forests.

Sound may be refracted. A large lens-shaped balloon of collodion film, when filled with carbon dioxide, acts as a condensing lens, for sound waves, and brings them to a focus by making the convex wave fronts concave. Beyond the focus the fronts again become convex outward, and the intensity of the sound diminishes.

Sound waves may be made to interfere. An illustration of this fact is the peculiar throbbing or beating effect produced when two tuning forks of slightly different frequency are sounded together.

107. Sound Waves Longitudinal. If a bell be struck or a tuning fork be put in motion, the layers of air immediately surrounding the vibratory system are thrown into vibrations of the same period, amplitude and form as the vibrations of the sounding body. This means that when the prongs of the tuning fork or the sides of the bell swing *outward* the layers of air are crowded together, producing a *condensation.* Since in a gas all the particles are free to move, this condensation is promptly handed on to the next layer of air particles, and by them to the next, and so on, each layer swinging *outward* in the same manner as the sounding body, only *a little later in time.* On the *inward* swing of the sounding body the layers of air immediately surrounding the body *are rarefied* by the receding body, and the pressure of the air from without promptly fills the rarefied space, only to produce a *rarefaction* where the *adjacent layers had been.* In this way the rarefaction is handed on from point to point, as the counterpart of the *condensation* that preceded it.

On the next *outward* swing of the body a second condensation is produced followed by a second rarefaction, and the phenomenon continues while the motion lasts. Now since the condensation and rarefaction occupy one half a vibration each in their production, it follows that *they together constitute the two parts of the sound wave,* and that the distance from one condensation or rarefaction to the *next* condensation or rarefaction is a *wave length,* λ, of the sound in question. Also in the condensation the particles are crowded together, and in the rarefaction they are drawn apart while the wave runs its own length during the time of a single vibration of the particles. Since these vibrations are *in* the line of propagation of the disturbance, it follows that *sound waves are longitudinal.*

108. Fundamental Differences. Loudness of Sound. Two sounds as perceived by the ear may differ in three respects. They may differ in *intensity,* in *pitch* or in *quality.* These three fundamental differences in sense perception owe their existence to fundamental differences in the vibratory move-

ments producing the sensations. It has already been noted that two waves may differ from each other in *amplitude*, in *wave length*, and in *wave form*. It will be shown hereafter (Arts. 108, 123, 141) that these three fundamental differences in wave motion determine the three fundamental characteristics of sounds as perceived by the ear.

Loudness or *intensity* of sound denotes in rough measure the kinetic energy of the vibrating medium as reported by the ear. Obviously in so complex a phenomenon as sense perception no rigid measure of the absolute intensity of a sound is possible. Not only does the same sound appear of different intensity to two different observers, but to the same observer on different days and in different states of health and nerve tension. Comparisons of the relative intensity of two or more sounds are possible, however, and show that the intensity of sound depends *upon the amplitude of vibration of the sounding medium.* We have seen that two waves of the same frequency may differ in the relative excursions executed by the vibrating particles. From elementary considerations it may be seen that if the excursions of the particle be increased, the average velocity of the particle must be proportionally increased, since the vibrations in each case are executed *in the same time.* The velocity of the vibrating particle is therefore proportional to the amplitude. But the kinetic energy of the particle is proportional to the *square of the velocity*, and consequently to the *square of the amplitude.*

The amplitude of vibration at any point of the sounding medium depends greatly upon the conditions under which the sound is produced. Considering the source of a sound and its location, the intensity of sound depends essentially upon three things.

(*a*) *Upon the amplitude of vibration of the sounding body.* Under otherwise identical conditions the relative displacement of the particles of the medium through which the sound is transmitted must increase with the amplitude of vibration of the sounding body, resulting, as was shown above, in an increase in the intensity of the sound at every point of the medium.

(*b*) *Upon the area of the sounding body.* The effect of area in modifying the loudness of a sound produced by a given body is essentially a question of transference of energy. A tuning fork, when once set in motion, possesses a definite amount of kinetic energy. If the fork be held in the hand, the sound produced is faint, but continues for several seconds. If the stem of the fork be placed firmly upon the top of a table, the note becomes loud and strong and soon dies away. In the latter case the top of the table is thrown into co-vibration with the fork, and the volume of air that is energized by this means is greatly increased. The store of energy is spent in the production of sound in a much shorter time.

(*c*) *Upon the distance from the sounding body.* The intensity of sound varies inversely as the square of the distance from the sounding body. This is known as "the law of inverse squares," and may be demonstrated as follows: Suppose a single sharp sound has been produced at a point O, as a center. This represents a definite amount of energy. Now at any instant this energy is resident on a spherical shell of radius r, and of area $4\pi r^2$ cm^2. If we call the intensity of sound I, the kinetic energy per unit area of this shell, then the total energy is $4\pi r^2 I$. An instant later this energy has been transferred to a shell of radius r_1, on which the intensity is I_1, hence the total energy on the shell is $4\pi r_1{}^2 I_1$.

Therefore $$4\pi r^2 I = 4\pi r_1{}^2 I_1 \tag{194}$$

whence $$\frac{I}{I_1} = \frac{r_1{}^2}{r^2} \tag{195}$$

or the intensity of sound varies inversely as the square of the distance from the sounding body.

From a physical standpoint the intensity of a sound is best defined as the *quantity of energy passing in unit time through unit area of a surface placed at right angles to the direction of the sound.* From this definition accurate measurements of the intensity are possible, but their realization is difficult and involves considerations beyond the limits of an elementary text.

CHAPTER XII

VELOCITY OF SOUND

109. Experimental Determinations. The problem of determining experimentally the velocity of sound in air presents peculiar difficulties. The nature of the problem demands that the measurement be made in free, open air, where the disturbing effects of winds, which vary both in magnitude and direction, and of local differences in temperature, are unavoidable and beyond the control of the observer. To these are added the errors of observation attendant upon measurements depending upon sense perception.

If the measurements are made by one observer who notes and records the time of seeing a flash, and of hearing the report of a distant gun, then each observation is affected by the errors due to the reaction periods of both sight and hearing. By this is meant the time required for any observer to see and record the flash, and to hear and record the sound. The reaction period for sight is different from that for hearing. The error for each operator is termed his "personal equation." Not only do these errors vary with different operators, but they vary with the varying conditions of health of the same operator on different days, and with the intensity of the sound. No allowance is made for the time occupied in transmitting the flash, since for small distances this time is practically zero.

A study of the various errors just mentioned has influenced the selection of methods of observation. If cannon be used as the source of sound, the method of "reciprocal firing," whereby the sound is transmitted first in one direction and then in the other between two stations, has been used to eliminate in large measure the errors due to the wind. This, however, requires

two observers, and introduces errors arising from differences in "personal equation" which are difficult to determine and cannot in general be wholly eliminated by exchange of observers. In the experiments of Stone in 1871, two operators, about three miles apart, recorded the time of hearing the report of a cannon placed some 640 ft distant from the first operator. On exchanging operators the error from "personal equation" was reduced to 0.02 sec.

The mean of Stone's results, reduced to 0° C, was 1090.6 ft per second, or 332.4 m per second.

***110. Experiments of Regnault.** Between the years 1862 and 1866 Regnault carried on an exhaustive series of experiments for the determination of the velocity of sound, both in the open air and in the water and gas pipes of Paris. In his researches, Regnault made use of an automatic recording apparatus, by means of which an electric current was broken at the instant of firing the gun, and the interruption of the current was recorded upon a smoked paper carried upon the drum of a chronograph. At the receiving station the sound wave entered a wide cone, at the smaller end of which it impinged upon a thin rubber membrane, and setting it in motion broke a second electric current, and so completed the record upon the cylinder of the chronograph. By this means, it would seem that the difficulties of personal equation were entirely obviated, but it was found that the membrane itself required time to receive and record the sound wave. The motion of the air particles cannot be imparted to the membrane instantly, and so a delay is caused in making the record, which is not constant, but increases as the sound grows more faint. Regnault made experiments to determine the amount of this error, and allowed for it in his computations.

In his experiments upon the velocity of sound in tubes, Regnault arrived at the following conclusions:

(*a*) In cylindrical pipes the intensity of the sound wave decreases with the distance, and more rapidly in small tubes than in large ones.

(*b*) The velocity of sound diminishes with the intensity. Loud sounds travel faster in tubes than faint ones.

(*c*) The velocity of sound in pipes increases as the diameter of the pipe increases, tending toward a limit in very wide tubes.

(*d*) The velocity is independent of the pressure, and of the mode of production of the sounds.

Regnault gave as the result of his investigations, after all corrections had been applied, the value for a faint sound in a very wide tube at 0° C

$$V_0 = 330.6 \frac{m}{sec}$$

More accurate determinations of this important constant, made by Violle and Vautier, gave for the velocity of sound in air at 0° C the value

$$V = 331.36 \frac{m}{sec}$$

and this value is now generally accepted as the best result yet obtained.

111. Theoretical Velocity of Sound. From purely theoretical considerations concerning the condition of the medium through which the sound wave passes, Newton, in 1686, developed a formula for the velocity of sound,

$$V = \sqrt{\frac{e}{d}} \tag{196}$$

where d is the density of the medium and e its coefficient of volume elasticity. This formula applies directly in the case of liquids and solids, while for gases it leads to erroneous results, unless there be applied certain corrections, the nature of which will appear later.

In the application of this formula to gases, it is to be shown that the coefficient of volume elasticity e is, for ordinary sounds, equal to the pressure to which the gas is subjected. Thus, let P, V and d represent the pressure, volume and density, respectively, of a given mass of air. Let the pressure be increased by a small increment dp, and as a result of this increase, let the

volume be diminished by a small amount dv, *the temperature remaining constant.* Then, by Boyle's law, we have

$$\frac{P}{P+dp}=\frac{V-dv}{V} \tag{197}$$

or

$$\frac{dp}{P+dp}=\frac{dv}{V}$$

whence

$$\frac{dp}{\frac{dv}{V}}=P+dp \tag{198}$$

but the left-hand member of this equation is by definition the coefficient of volume elasticity, e (Art. 58), hence $P+dp=e$.

For ordinary sounds the change in pressure dp, produced in the condensation or rarefaction, is negligible as compared with the atmospheric pressure, and hence may be disregarded. Under this assumption our formula becomes

$$V=\sqrt{\frac{P}{d}} \tag{199}$$

where P is the reigning barometric pressure, and d the corresponding density of the air.

This would indicate that the velocity of sound is independent of the intensity, and this assumption seems to be sustained by observation in the case of all sounds where the change of pressure dp, produced by the condensation, is negligible in comparison with the barometric pressure, P. In disturbances where dp is not negligible, the velocity is represented by the formula

$$V=\sqrt{\frac{P+dp}{d}} \tag{200}$$

and increases as the intensity increases. In loud sounds, as those produced by explosions or the discharge of cannon, the change in pressure dp may even exceed P. Captain Parry, who made a series of measurements upon the velocity of sound in the arctic regions, relates that the report of a cannon was fre-

quently heard by a distant observer, before the command to fire was heard. Other observers have repeatedly found that very loud sounds travel more rapidly than those of ordinary intensity.

*** 112. Application of Newton's Formula.** It has been said that Newton's formula

$$V=\sqrt{\frac{P}{d}}$$

does not apply directly to the case of gases. If the values of P and d be substituted in this formula, we have for standard conditions, $P = 1{,}012{,}630$ dynes per square centimeter, and $d = 0.001293$ g per cubic centimeter. Whence

$$V=\sqrt{\frac{1{,}012{,}630}{0.001293}}=27{,}927\frac{cm}{sec}$$

This is only 84 per cent of the velocity of sound as determined by experiment. The discrepancy between the theoretical and observed value was recognized by Newton, who sought to account for the difference by means of a number of ingenious suppositions, none of which, however, were justified.

113. Laplace's Correction. It was not until 1816[1] that the error in Newton's formula was pointed out and corrected by Laplace. It was shown (Art. 111) that for small disturbances the coefficient of volume elasticity e, in the formula

$$V=\sqrt{\frac{e}{d}}$$

could be replaced by the pressure P, of the gas.

This deduction was made upon the assumption that Boyle's law would hold for the phenomena. This would mean that the air in which the condensations and rarefactions of the sound wave were produced should remain throughout *at a constant temperature;* or, in other words, that the condensations and rarefactions were to occur under *isothermal conditions.*

Laplace pointed out that in the case of a sound wave, where

[1] *Ann. Chim. et Phys.* 3, p. 238, 1816.

the compressions and rarefactions follow each other at the rate of several hundred per second, isothermal conditions do not hold, since there can be no time for the air to assume a uniform temperature. The coefficient of elasticity of the air, that really does enter the equation for the velocity of sound, is therefore to be determined, under the condition that *no heat is to escape from the gas when it is compressed, and none is to come to it when it is rarefied.* Such conditions for a gas are called *adiabatic conditions*, and the coefficient of elasticity of a gas obtained under such conditions is called *the coefficient of adiabatic elasticity* (Art. 184).

Now it is well known that if a mass of gas be compressed suddenly, its temperature is raised, and the elastic tension of the gas is greater *because of this increase in temperature.* Careful experiment has shown that for air, the coefficient of *adiabatic elasticity* is 1.41 times the coefficient of *isothermal elasticity.* Hence the pressure P must be multiplied by 1.41 to represent the facts. The corrected formula is therefore

$$V=\sqrt{\frac{1.41\,P}{d}} \tag{201}$$

When the correcting factor is introduced into the computations in the previous article we have

$$V=\sqrt{\frac{1.41\times 1{,}012{,}630}{0.001293}}=332.3\ \frac{m}{sec}$$

which agrees admirably with the value found by experiment.

114. Correction for Temperature. Since the air is free to expand, an increase in temperature of the air as a whole will affect the *density* of the air, but will leave the pressure unchanged. If α be the coefficient of expansion for gases, then any bulk of air at 0° C will have a bulk $(1+\alpha t)$ times as large at $t°$ C. Since the densities are inversely as the volumes, then d_t, the density of the air at $t°$ C, will be given by the equation

$$d_t=\frac{d}{1+\alpha t} \tag{202}$$

Hence the velocity of sound at $t°$ C will be

$$V_t = \sqrt{\frac{1.41 \times P}{d_t}}$$

or

$$V_t = \sqrt{\frac{1.41 \times P(1+\alpha t)}{d}} \tag{203}$$

If V_0 be the velocity of sound at 0° C, then

$$V_t = V_0\sqrt{1+\alpha t} \tag{204}$$

If we set $t = 1°$ C and $\alpha = 0.003665$, then the *increase* in velocity for a rise in temperature of 1° C is

$$\begin{aligned} V_1 - V_0 &= 331.36\sqrt{1+0.003665} - 331.36 \\ &= 0.00183 \times 331.36 \\ &= 0.606\,\frac{m}{sec} \text{ or } 23.8\,\frac{in}{sec} \end{aligned}$$

Hence the velocity of sound increases about 60.6 cm per second, or 23.8 in per second for an increase of 1° C.

115. Velocity of Sound in Solids and Liquids. For solids and liquids Newton's formula is applicable at once. For copper the coefficient of elasticity e is about 12×10^{11} dynes per square centimeter and $d = 8.8$ g per cubic centimeter, therefore

$$V = \sqrt{\frac{12 \times 10^{11}}{8.8}} = 369{,}300\,\frac{cm}{sec} = 3693\,\frac{m}{sec}$$

For water the density is 1 gram per cubic centimeter and the *compression for an increase in pressure of one atmosphere is* 0.0000499; hence

$$V = \sqrt{\frac{1{,}012{,}630}{1 \times 0.0000499}} = 142{,}500\,\frac{cm}{sec}$$

$$= 1425\,\frac{m}{sec}$$

The best experimental determinations of the velocity of sound in water give a mean result of 1435 m per second.

CHAPTER XIII

REFLECTION AND SUPERPOSITION OF SOUND WAVES

116. Huygens's Principle. Suppose a sound wave originate at a, as a center of disturbance. The wave will travel outward in space in all directions in the form of a spherical shell. Let Fig. 58 represent the trace on the plane of the paper of a section through this wave shell. At the end of a certain time, the wave will have reached the position *mcn*. This surface, marking the location of all points in the same phase of vibration, is by definition a *wave front*. At an instant later it will have taken the position $m'dn'$, and all points in this new wave front will be vibrating with the same motion possessed *now* by the points on the wave front *mcn*. The disturbance at the center has thus reached any point in *mcn* by the disturbance of all points in the medium through which it has passed. Hence the subsequent disturbance at any point outside the wave *mcn* is to be regarded as the resultant effect, produced at that point, of all the wave disturbances originating in the individual points of the wave front *mcn* as centers.

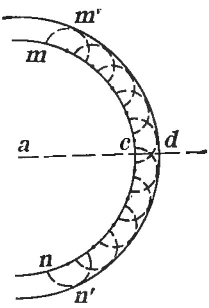


FIG. 58.

According to *Huygens* we are to consider *every point on a wave front as a new center of disturbance, from which waves are propagated outward as from the original center, and the effect at any point external to the wave is to be regarded as the resultant of the combined action of all these elementary waves.*

If we take all the points on the wave front *mcn* as centers and about them describe small circles representing the limits reached by the elementary waves in a very short space of time, then it will be seen that these little waves unite to form a *new*

wave front, which is a common tangent or *envelope* to them all. This is to be regarded as the new wave front. Professor Stokes has shown that the elementary waves destroy each other except at the surface of their common envelope.

117. Reflection of Sound. Sound waves are reflected regularly from any smooth surface of different density from that of the medium in which they are traveling, and the sound seems to come from some point behind the reflecting surface. By a "smooth" surface is to be understood a surface whose inequalities are small in comparison to the wave lengths to be reflected. Thus a brick wall, a hillside or even a row of trees may present a surface that is relatively smooth to the waves of sound, while for the reflection of light waves only the most highly polished surfaces are "smooth."

Tyndall showed that sound waves are reflected not only from solids and liquids but from layers of gas of different density as well. Thus a flame from a fishtail gas burner and the heated air above it reflect sound waves to a marked degree. In the same way ascending currents of warm air, over parts of the earth unequally heated by the sun, serve to reflect and scatter sound waves, as do also ascending layers of air loaded with water vapor.

In some cases the nature of the sound is modified by the nature of the reflecting surface, as in the case of a sharp, clear sound reflected by the separate bars of a picket fence. Here the separate pickets act as independent sources of sound waves and the short, sharp sound comes back as a clear musical note of definite pitch.

Echoes are produced when the reflecting surface is at a distance of more than 16.5 m. The ear cannot separate syllables occurring more rapidly than about ten per second. Hence a sound and its echo will be heard separated if the distance to the reflecting surface is greater than half the distance which sound would travel in one tenth of a second.

Applications of the principle of reflection of sound are seen in the speaking trumpet or megaphone, the ear trumpet, the stethoscope and in the curved surface of the external ear which

acts as a reflector to concentrate the sound waves into the auditory canal.

118. Reflection at End of Cylindrical Pipe. A case of special interest is found in the reflection of a sound wave at the end of a tube or pipe in which it is traveling. If the reflections occur in a closed pipe, against the solid end of the pipe, the direction of the wave is reversed, but the character of the disturbance is not reversed, a condensation is reflected as a condensation and a rarefaction as a rarefaction. This is expressed by saying that at the closed end of a pipe the sound wave is reflected *with change of sign in the velocity of the air particles, but without change of sign in the condensation.*

At the end of an open pipe, however, the result is different. When the condensation reaches the open end of the pipe, the last layer of air, being no longer hemmed in by the walls of the tube, expands more freely than the layers in the tube and produces a rarefaction which is greater than that of the layers inside the tube. A new disturbance thus enters the tube at the open end, in the nature of a rarefaction, and is propagated backward, while the original condensation pursues its course in the free air. This is expressed by saying that reflection occurs at the open end of a pipe, *without change of sign in the velocity of the air particles, but with change of sign in the condensation.* At the open end of a pipe, therefore, a condensation is reflected as a rarefaction, and *vice versa.* This principle will be found of great importance in the theory of open and closed organ pipes.

119. Superposition of Sound Waves. If two stones be dropped a small distance apart into still water, each is seen to become the center of a system of circular waves, widening in all directions as they run. If the two sets of waves cross each other, they are seen to pursue their own way in each case as if there were no other wave on the surface of the water. In general, when two systems of waves traverse a medium simultaneously, while the elevation or depression of either alone relative to the disturbed surface is the same as it would have occasioned if the other were not there, yet the total displacement of the surface

relative to the undisturbed level is at any point and at any instant the algebraic sum of the disturbances which each of the systems would produce separately.

If an elevation of the first system be superposed upon an equal elevation of the second system, the total height of the water above the original level will be double that of one of the two waves alone. If an elevation of one system meet an equal depression of the other system, the original level is not changed.

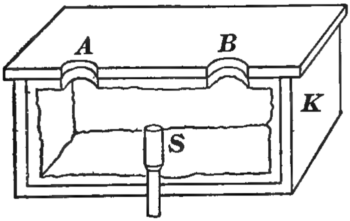


FIG. 59.

In a similar way, sound waves, when coexistent in the same medium, form at every instant composite waves, made up of the separate systems. This superposition of wave systems is termed *interference*.

120. Principle of Interference. Suppose a small whistle to be placed in a box, heavily padded within to prevent any transmission of sound waves through its walls, and let there be in the top of the box two identically similar openings, *A* and *B* (Fig. 59), symmetrically situated with respect to the whistle. We have here two identical sources of sound, from which condensations and rarefactions proceed outward in all directions at the same instant, and such that air particles, equidistant from the two sources, on being disturbed by the individual sources *A* and *B*, have their motions outward *from* or inward *toward* their respective openings *at the same instant.* If now we consider a plane PP' (Fig. 60), normal to AB at its middle point, it is clear that two similar disturbances, starting out from *A* and *B* respectively, at any instant, will reach any point in PP' in the same time, since the paths are exactly equal. It is also clear that each wave

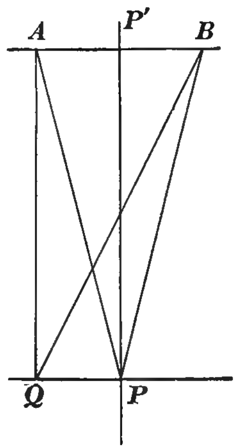


FIG. 60.

will, at any point, have a component of its motion normal to this plane, PP', and that these normal components will, at one instant, both be directed *toward* the plane, and a half period later they will both be directed *away* from the plane. In this plane, therefore, we have at every point condensation meeting condensation, and rarefaction meeting rarefaction, thus producing *increased condensations and increased rarefactions, in this plane* PP'. This plane is therefore a locus of maximum *change in density*, and is consequently a locus of *maximum intensity of sound as perceived by the ear.*

It is also to be noted that two air particles symmetrically situated on either side of this plane, although moving toward each other, actually meet *in the same phase*, since their displacements from their respective positions of rest are equal and have been produced *in the same time.*

FIG. 61.

In any other plane drawn through AB, the *condensations* and *rarefactions*, and consequently the *intensity*, will present a series of fixed maxima and minima. If we consider the disturbance from the single point O (Fig. 61), along two lines OM and OM' very near to each other, so that we may assume that the motions have experienced the same conditions in passing from O to M and M', then we shall find the particles in M and M' to be in the *same phase of vibration* if the difference of the distances OM and OM' be an even number of half wave lengths. They will be in *opposite phase* if the difference of the paths OM and OM' be an *odd* number of half wave lengths. In the first case the velocity and displacement of the two particles at M and M', *due to the single source at* O, will be equal and in the same direction; in the second, they will be equal and in opposite directions.

From the foregoing considerations we see that when sound waves from two identical sources A and B (Fig. 60) meet at a point Q, we shall have *maximum* change in *density* if

$$AQ - BQ = 2n\lambda/2 \tag{205}$$

and *minimum* change in *density* if

$$AQ - BQ = (2n + 1)\lambda/2 \tag{206}$$

that is, according as the difference of path $AQ - BQ$ is an *even* or an *odd* multiple of $\lambda/2$.

Hence for maximum intensity the difference in path is an *even* number of half wave lengths; for minimum intensity the difference in path must equal an *odd* number of half wave lengths. In this way two equal sounds may be added together so that at certain points in space they will mutually reënforce each other, and the sound will be very loud; in other points where the vibrations meet in opposite phase, one sound added to another may produce a sound of greatly diminished intensity.

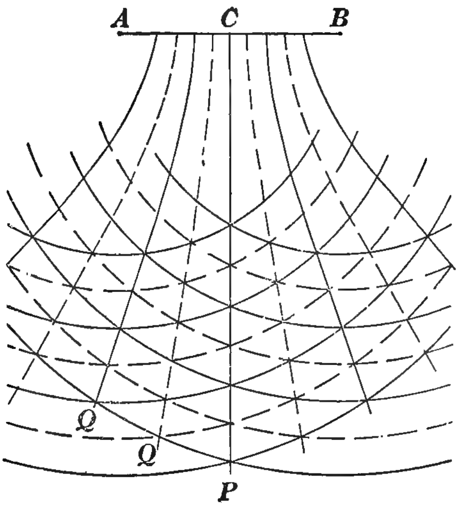


Fig. 62.

121. Curves of Maximum and Minimum Intensity. The interference of waves from two identical sources of sound may be shown graphically by the following construction: About the point sources A and B (Fig. 62) describe arcs of circles with equal radii, where the full line arcs denote condensations and the dotted arcs represent rarefactions. The shortest distance between any two concentric arcs *of the same kind* is therefore some multiple of the distance from one condensation to another condensation, or from one rarefaction to another, *i.e.* some multiple *of a wave length* λ, of the sound in question. Similarly, the distance between a *full line arc* and the *next adjacent dotted arc* described about the same center is *a half wave length.*

From the foregoing considerations it is evident that at all

points on the line CP, normal to AB at its middle point, the two sets of waves *meet in the same phase*, condensation from A meeting condensation from B, and rarefaction from A meeting rarefaction from B, as indicated by the intersection of the corresponding arcs. The full line CP is therefore *a line of maximum intensity*, and the remaining full lines cutting AB, being also loci of particles in the same phase of vibration, are *also lines of maximum intensity.* The dotted lines on either side of CP, drawn through the intersections of rarefactions from A with condensations from B, or *vice versa*, are consequently lines of *minimum intensity of sound.* These curves are all hyperbolae, defined by the equation

$$AQ - BQ = \pm n\lambda/2$$

where Q is any point whatever, and n is any integer 0, 1, 2, 3, etc. The lines of maximum intensity are given by the even values of n, and those of minimum intensity by the odd values.

We should remember that there is *no loss of energy* due to interference. At the maxima the amplitude is *double*, and the energy *four times* that due to a single source. At the minima, the amplitude and the energy are both zero. Hence the average energy over the surface is *twice* that due to a single source, as it should be.

122. Experiments illustrating Interference.

(*a*) *Sources Identical*

If a tuning fork be sounded and rotated slowly before the ear, an intermittent or pulsating sound will be heard. At four certain positions, it will be found that the sound of the fork seems to disappear almost entirely, only to reappear again in force on moving the fork. In this case the two identical sources are the two prongs of the fork, which are vibrating in opposite directions as indicated in the arrows (Fig. 63). As the prongs approach, there is a rarefaction on the outside of each prong at f and g, and a double condensation on the inside, starting out toward d and e. These two sets of disturbances

differing in phase by half a period meet along the dotted lines and produce a minimum sound at all points in these lines. Outside these lines the sound may be heard as usual.

To show that this is a true case of interference the fork may be held over a resonator which responds loudly, and rotated till the sound falls to a minimum. On slipping a small cylinder of paper over one prong of the fork, so as to cut off one set of waves, the sound reappears, but disappears again if the paper be removed.

If a large fork or organ pipe be sounded in a large room and an observer walk about the room, certain places will be found where the sound is uncomfortably loud while at others almost no sound is heard. In this case interference occurs between the direct waves from the fork and the systems of waves reflected from the sides of the room.

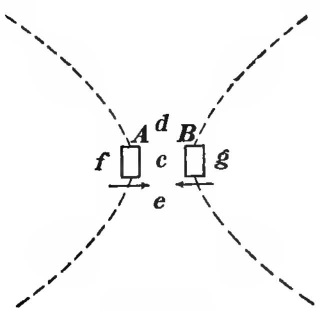


FIG. 63.

(*b*) *Sources not Identical.* — ***Beats***

If two tuning forks, tuned to unison and furnished with resonance cases, be sounded together, a full even tone is heard. If now the prongs of one fork be weighted with wax, so as to decrease slightly the frequency of that fork, we shall find that when sounded together, the forks give out a throbbing or beating tone. If the pellets of wax be made larger, the number of beats per second increases.

In this case the phase difference $\lambda/2$ is not due to the difference in path traversed by the waves from the sources, as hitherto, but has resulted from a difference in the *frequencies* of the two forks. If the forks make respectively 100 and 101 vibrations per second, then the second fork gains one vibration upon the first every second. Under these conditions, if the two forks *started* together in the same phase, they would be in opposite phase at the end of half a second, and in coincidence of phase again in another half second. From this it appears that two

forks whose frequencies are m and n vibrations per second will make $m - n$ beats per second.

Two organ pipes of the same pitch, when mounted upon a wind chest, will give loud beats if, while they are sounded, a card be slipped slightly over the lip or the end of one of the pipes. A tuning fork mounted upon a resonance case and sounded gives distinct beats if carried rapidly *toward* or *away* from a reflecting wall. The experiment may be rendered more striking by swinging the fork as a pendulum, while sounding, at a short distance from the wall.

MUSICAL RELATIONS

CHAPTER XIV

MUSICAL SCALES

123. Pitch. The pitch of a sound depends upon the vibration frequency of the sounding body. When the number of vibrations per second is great, the pitch of the tone is high or *acute*; when small, the pitch is low or *grave*. If two sounds are produced by the same number of vibrations per second, they are said to have the same pitch, or if sounded together, they are said to be in *unison*. *Musical sounds* are those which produce a pleasing effect upon the ear and *have a definite pitch*. A noise is a confused mass of sonorous vibration in which the ear is unable to detect any definite pitch. In a musical sense, the pitch of a sound may also refer to the relative position of the sound upon some arbitrary scale of reference adopted by musicians.

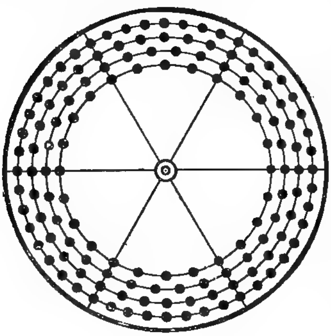


FIG. 64.

The pitch of a sound may be determined by means of an instrument called the *siren* (Fig. 64). The siren consists of a pasteboard or metal disk, bearing on its circumference a series of concentric circular rows of equidistant holes about 3 mm in diameter, and mounted on an axis which can be rapidly revolved in front of a nozzle delivering air from a blower. When an opening comes in front of the nozzle the air rushes through, forming a condensation, followed by a rarefaction during the

interval in which the air is cut off owing to the momentum of the air particles. In this way is formed a series of regular puffs, which gradually blend into a low musical tone, whose pitch rises as the speed of rotation of the disk increases. The siren tone may be tuned to that of the given note, and its frequency determined from the angular velocity of the disk and the number of holes passing the nozzle in one revolution.

124. The Diatonic Scale. The rule for consonant intervals extends to combinations of several sounds. In order that three or more tones when sounded together may be concordant, it is necessary that their respective intervals not only with the fundamental, but also with each other, should be expressed by simple ratios. Thus when we sound together three notes whose frequencies are as 4 : 5 : 6, there is produced a pleasing effect. This combination of three tones is called a *major triad.* The diatonic scale is built upon three sets of such triads.

The notes of the scale are indicated by the letters C, D, E, F, G, A, B, c. These letters may represent the frequencies of the various notes as well.

In the key of C, the three major triads are

$$\left.\begin{array}{ll} \textit{Tonic} & C : E : G \\ \textit{Dominant} & G : B : d \\ \textit{Subdominant} & F : A : c \end{array}\right\} \; 4 : 5 : 6$$

The vibration ratio in terms of C, the fundamental, are obtained as follows:

$$\frac{E}{C} = \frac{5}{4} \quad \text{or} \quad E = \frac{5}{4}C$$

$$\frac{G}{C} = \frac{6}{4} \quad \text{or} \quad G = \frac{3}{2}C$$

$$\frac{B}{G} = \frac{5}{4} \quad \text{or} \quad B = \frac{5}{4}G = \frac{5}{4} \cdot \frac{3}{2}C = \frac{15}{8}C$$

In this way, the frequencies of the entire seven notes may be expressed in terms of the fundamental C, the octave c being set

equal to 2 *C*, and we have the following relations between the various notes of the scale:

Frequency	72	81	90	96	108	120	135	144
Name	*C*	*D*	*E*	*F*	*G*	*A*	*B*	*c*
Syllable	*Do*	*Re*	*Mi*	*Fa*	*Sol*	*La*	*Si*	*Do*
Ratio	$\frac{1}{1}$	$\frac{9}{8}$	$\frac{5}{4}$	$\frac{4}{3}$	$\frac{3}{2}$	$\frac{5}{3}$	$\frac{15}{8}$	2
Intervals		$\frac{9}{8}$	$\frac{10}{9}$	$\frac{16}{15}$	$\frac{9}{8}$	$\frac{10}{9}$	$\frac{9}{8}$	$\frac{16}{15}$

The fractions termed *intervals* are obtained by dividing each ratio by that of the note immediately below it.

From this it appears that in the perfect diatonic scale there are three different intervals, 9/8, 10/9, 16/15. The first two intervals are termed *whole tones*, and the last a *half tone.*

The *minor triad* is composed of three notes, whose frequencies are 10 : 12 : 15. and a *minor scale* similar to the *major scale* may be built upon three of such triads. The vibration ratios for the various notes in terms of the fundamental may be obtained in the same way. On making the computations, we have

Frequency	72	81	86.4	96	108	115.2	129.6	144
Name	*C*	*D*	*E*	*F*	*G*	*A*	*B*	*c*
Ratio	$\frac{1}{1}$	$\frac{9}{8}$	$\frac{6}{5}$	$\frac{4}{3}$	$\frac{3}{2}$	$\frac{8}{5}$	$\frac{9}{5}$	2
Intervals		$\frac{9}{8}$	$\frac{16}{15}$	$\frac{10}{9}$	$\frac{9}{8}$	$\frac{16}{15}$	$\frac{9}{8}$	$\frac{10}{9}$

It will be found that three additional notes will be needed, viz. three notes below *E*, *B* and *A*, in order to produce the minor scale.

125. Musical Intervals. When two tones are sounded together, the ear recognizes a certain relationship, or want of relationship, between them, dependent upon their relative pitch, but entirely aside from their absolute vibration frequencies. This relationship is termed a *musical interval*, and is expressed

as a simple ratio between the vibration frequencies of the two tones in question.

A number of these ratios have specific names in musical nomenclature, arising for the most part from the number of the note in the series. The interval between two notes whose vibration frequencies are in the ratio 1/1 is called *unison;* 2/1, an *octave;* 3/1, a *twelfth;* 4/1, the *double octave;* 3/2, a *fifth;* 4/3, a *fourth;* 5/4, a *major third;* 6/5, a *minor third;* 5/3, a *major sixth;* 8/5, a *minor sixth.*

This completes the list of so-called *consonant intervals*, although the list may be and probably will be extended in the course of time. It is interesting to note that the third, both major and minor, were originally classed among the dissonant intervals, and the minor third was not regularly used until the middle of the eighteenth century.

***126. Transposition.** In order to accommodate different voices or instruments, it is frequently desirable to change the keynote of the scale from *C* to some other note in the scale. The vibration ratios would then have to be applied to the *new keynote* as a fundamental, and the corresponding frequencies for the several notes computed. If it were desired to begin the scale with *D*, then, on computing the frequencies for the scale, it would be found that, beside the keynote *D* and its octave *d*, the *G* and *B* were right, and that the *A* and *E* differed but slightly from the required frequency, but the notes *F* and *c* would be found to be too low in each case. This must be remedied by the introduction of two new notes, *F* sharp and *c* sharp, in order to sing or play in the key of *D*. These two sharps are introduced at the beginning of the staff, and form the signature of the key.

***127. The Tempered Scale.** Since each change of key entails the introduction of new notes, both for major and minor scales, it is apparent that the number of notes demanded for each octave, in order to render a piece of music in any key, would be very greatly increased, so much so, indeed, that in the case of an instrument of fixed tones, as the piano or organ, it becomes

practically impossible to manipulate so many keys. On this account, a compromise system, known as the system of *equal temperament*, has been adopted. In this system the whole tones are made all alike, and the half tones are half the whole tones. In other words, *there is an equal interval between each pair of consecutive notes.* There are thus added five new tones to the octave, making thirteen tones in all. The common ratio between the frequencies of any tone and the tone next above it is the *twelfth root of 2*, or 1.059. In any instrument tuned to this system, the only accurate intervals are the octaves, all the others being slightly false. The fifths are slightly flat, and the thirds are too sharp. Music rendered in this system is considered inferior to that played in just intonation. Trained singers, and performers upon instruments like the violin or slide trombone are free from the limitations of the system of equal temperament, and in many cases approximate closely the intonation of the diatonic scale.

RESONANCE PHENOMENA

CHAPTER XV

VIBRATORY PHENOMENA AND RESONANCE

128. Composition of Vibrations at Right Angles. In the study of vibratory motion, some curious and beautiful results are obtained from combining two simple harmonic motions at right angles to each other. Owing to the rapid motion of sounding bodies the eye is unable to follow them, and some special device is necessary. In Fig. 65 are shown two tuning forks *L* and *M*, each furnished with plane mirrors, one set to vibrate vertically, the other horizontally. If now a beam of light from *R* be allowed to fall upon the mirrors successively and be reflected to the screen *S*, when the fork *L* is set vibrating, the spot is seen drawn out into a vertical band. Similarly, if the

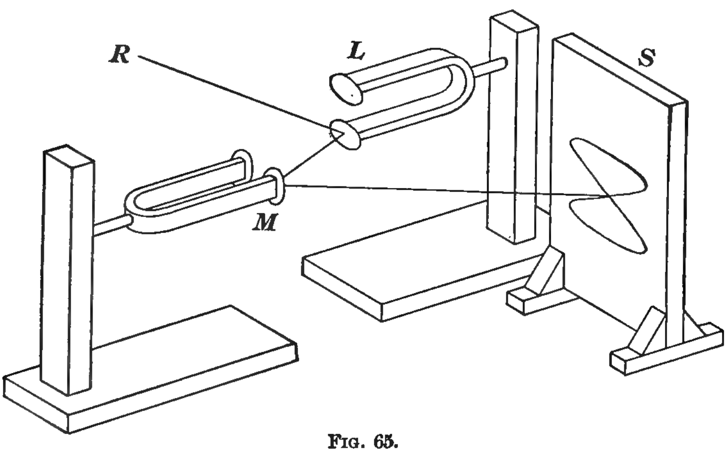


FIG. 65.

fork *M* be set in motion and *L* kept at rest, the spot is drawn out into a horizontal band. If the two forks be vibrated at the same time, the spot is made to follow the motion of the two vibrating systems and traces some form of what is known as a Lissajous' curve, of which various forms are shown in Fig. 66.

If the frequencies of the two forks be in the ratio of 1:1, then the characteristic figure will be an ellipse, having for its

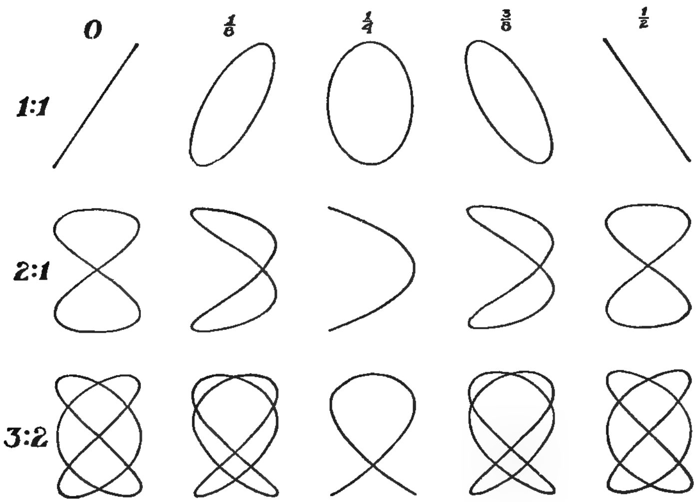


Fig. 66.

special forms the two straight lines. If the tuning of the forks be exact, the figure is motionless upon the screen and gradually decreases in size as the amplitude of the vibrations of the forks sinks to zero. In most cases, however, the tuning is only approximate, and the figure takes the successive forms indicated (Fig. 66), passing from left to right and back again. When the figure has run through the complete cycle, we know that one of the forks has gained or lost one complete vibration as compared with the other. We have thus a method of observing beats optically, and of determining the relative frequencies of vibrating bodies with great precision.

If, however, the frequency of fork M be twice that of fork L, we shall have the curve shown in Fig. 65, of which different forms appear in Fig. 66. In this figure are also shown the curves for the interval 3:2. The curves are drawn for the phase differences indicated at the top, where these differences are stated with respect to the component having the higher frequency.

These phase differences are exact at the beginning and close of any complete period. Thus in the ratio 3:2 the x motion has the higher frequency, 3, while the y motion has a frequency of 2. At the beginning the x motion leads in phase by 0, $\frac{1}{8}$, $\frac{1}{4}$, $\frac{3}{8}$ or $\frac{1}{2}$ T, as the case may be, and also when the motion on the x-axis has made *three* vibrations, and that on the y-axis has made *two*.

This experiment was first described by the French physicist Lissajous, in 1857, and the curves are known as Lissajous' curves. The method is applicable to the study of any vibrating system upon which a bright point as a minute globule of mercury can be fixed, while the fork with which the system is to be compared is armed with a lens of low power through which the mercury globule may be viewed by a microscope.

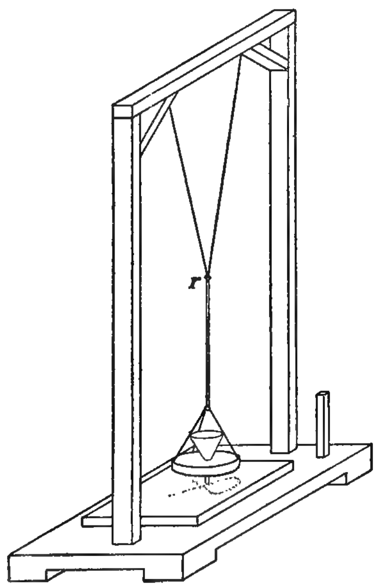


FIG. 67.

The same figures may be obtained by means of the Blackburn's pendulum, shown in Fig. 67. In this apparatus a heavy lead disk carries a funnel filled with sand, ink, or other material for leaving a trace of the motion upon a prepared paper beneath it. The disk is hung from two cords about one

meter in length, and over the two cords is slipped a small ring *r*, by means of which the system is divided into two pendulums of different period, hung from the same support. On setting the disk vibrating it traces the figures characteristic of the ratios represented by the square roots of the lengths of the entire pendulum and the part below the ring.

129. Graphical Method for Lissajous' Figures. Draw two concentric circles (Fig. 68), with radii proportional to the amplitudes of the two harmonic motions, and through their common center O draw the rectangular diameters AB, CD.

Divide each quadrant of both circles into the same number of equal parts; some multiple of four is usually most convenient. Through the points of division of the circle AB draw lines parallel to CD, and through the divisions of CD draw lines parallel to AB. The resulting rectangle will contain all the figures arising from any possible combination of two simple harmonic motions of commensurable periods; and the curves will, in general, be tangent to the sides of the rectangle.

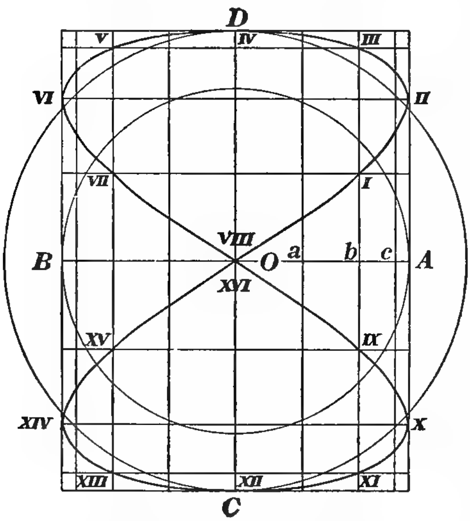


FIG. 68.

The center of the circles corresponds to a phase difference of zero between the two components, that is, to $\delta = 0$; and it is taken as the starting point for tracing all curves of phase difference zero or $T/2$.

If, as in Fig. 68, the circles have been divided into sixteen equal parts, then each point of intersection on the diameter AB corresponds to a phase difference of $T/16$; that is, to one sixteenth of a period. Hence if we start to trace a curve from a in the figure instead

of from O, we shall produce the curve corresponding to a phase difference of $T/16$. This means that at the instant when the y component passes through AB in a positive direction, and the y displacement is therefore zero, the x component has already reached a in the positive direction, or is in advance of the y component by Oa or $T/16$. In like manner, b corresponds to a phase difference of $T/8$, c to $3\,T/16$, and A to $T/4$. Returning toward O, it will be seen that c also corresponds to a difference of phase of $5\,T/16$, b to $3\,T/8$, a to $7\,T/16$, and O to $T/2$, with larger values for points to the left of O.

Suppose now that we wish to trace the curve corresponding to the vibration frequencies one to two; two for the horizontal and one for the vertical component, and with no difference of phase. Starting from O, we count two points horizontally to the right and one up and reach I; again two to the right and one up for point II, and so continue, numbering the points in order until we pass through the starting point in the same direction as at first, being careful always to complete the motion in one direction before beginning the retrograde motion. An excellent check upon the accuracy of the location of the points is found in the fact that points equidistant from the axis of symmetry AB differ in number by eight in every case; that is, by half a vibration.

If now a smooth curve be traced through the points in order, we see that the moving point, being subject to both motions, describes two spaces horizontally and one vertically in the same interval of time, and consequently passes through the corners of rectangles two spaces long and one space high in every case. The spaces themselves increase or decrease according to the simple harmonic law. Great diversity of figure may thus be obtained with successive differences of phase between the two component motions.

To combine two motions of frequencies two to three, we should simply count three spaces in one direction and two in the other and proceed in other respects as already described.

130. Free and Forced Vibrations. If a system when displaced from its position of equilibrium is urged to return to that posi-

tion by forces, either internal or external, which vary directly as the displacement, it will, when freed from the disturbing force, execute simple harmonic vibrations about its position of equilibrium as a center. The *period* of such a vibration *depends upon the nature of the system*, and is independent of either the amplitude of vibration, or of the forces tending to oppose the vibration, provided each be small. Such a vibration is called a *free vibration*, and the period is termed the *free or natural period* of the system.

In the ideal case of no friction, a body once started would continue to vibrate forever, since there would be nothing to stop it. In nature, however, all vibrations are checked with more or less promptness by means of opposing forces which may be designated under the general name of friction. In such cases the amplitude gradually decreases to zero, while the period remains constant, and is independent of the friction, provided it be small. Such a vibration is termed a *damped vibration*. Examples of *free vibration* are seen in the motion of a simple pendulum, of a guitar string or of a tuning fork when bowed and allowed to swing freely.

When a system that is free to vibrate is subjected to the action of a periodic force that varies as an harmonic function of the time, we have the conditions necessary for a *forced vibration*. The ensuing motion is the response of the system to the impressed, external force, and continues so long as the force continues. Examples of *forced* vibration are seen in the motion of the pendulum of a clock or the balance wheel of a watch, in the vibration of a tuning fork driven by an electric current, or of the sounding board of a piano or body of a violin. Since any free vibration is always more or less damped, and therefore soon sinks to zero, it follows that any *maintained* vibration is a *forced* vibration; the motion of the vibrating system being maintained by an external impressed force which varies with the time. The *period of the forced vibration* is the *period of the force*, and the amplitude is proportional to the force.

The characteristics of *free* and *forced* vibrations may be

contrasted as follows: A *free* vibration gradually *dies away* on account of frictional forces. A *forced* vibration *is maintained* so long as the impressed force continues, and when it ceases the free vibration ensues and gradually sinks to zero. The *period of a forced vibration* is the *period of the impressed force*, while the period of a *free vibration* depends upon the *constitution of the system*, and is entirely independent of the forces causing it, so long as the amplitude is small.

131. Resonance. A special case of *forced vibration* is that in which the period of the impressed force coincides with the free period of the system. In such a case the system rapidly absorbs energy from the individual, periodic pulses, and soon vibrates with large amplitude. Theoretically it is due to the friction alone that the amplitude does not become infinite. *It thus appears that a system free to execute vibrations of a definite period is capable of selecting and absorbing from the surrounding medium energy in the form of vibrations of the same period as those which it can execute.*

This is known as the *principle of resonance*, which finds its applications in every department of physics. Resonance depends upon the cumulative effect of small impulses applied to a system at exactly the proper time to produce the maximum effect.

132. Illustrations of Resonance. Two strings stretched upon a sonometer, if tuned to unison, will mutually transmit vibratory motion, by means of synchronous impulses sent through the air and through their common support. If either string be set in vibration, the other begins to vibrate.

Let two heavy pendulums of the same period be mounted upon a wooden frame which yields slightly to their motion, and let one be set vibrating while the other remains at rest. In a few minutes it will be seen that the second pendulum is acquiring vibratory motion through the support. Its motion gradually increases until the two are swinging with equal amplitude, but with a *phase difference of a quarter period.* The second pendulum continues to lag behind the first, gradually

absorbing its energy until the first is brought to rest, after which the phenomenon is repeated in the reverse order.

If a tuning fork be held over the mouth of a tall jar (Fig. 69) partly filled with water, it will be found, on pouring in more water, that for a certain length of air column the sound of the fork is powerfully reënforced. If the fork be removed and a blast of air from a flat tube be blown across the top of the jar, the sound produced will be in unison with that emitted by the fork. The cylinder thus behaves as a stopped organ pipe (Art. 140) and the air column is very nearly equal to one fourth the wave length of the sound produced by the fork. The hollow wooden cases used to support tuning forks are in reality closed or open pipes tuned to reënforce the tone of the fork.

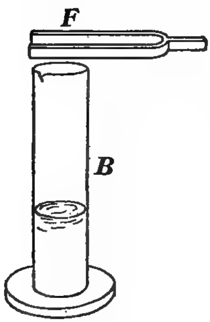


FIG. 69.

Let two tuning forks mounted upon suitable resonance cases and accurately tuned to the same pitch be placed at opposite ends of a room. If one be bowed and then quieted, it will be found that the other is sounding audibly. Accurate tuning is necessary for success in this experiment. If a number of forks of different pitch be sounded together, the second fork responds to none but the one of its own frequency. Obviously the fork can absorb from the air only those wave lengths of sound which it itself can emit.

A heavy bridge is often set to vibrating vigorously by the footfalls of a small dog trotting across it. Soldiers when crossing a bridge are commanded to break step to avoid the possibility of synchronous vibration of the bridge.

The resonators of von Helmholtz (Fig. 70) consist of hollow spheres of brass, furnished with a tubular opening for the reception of the sound wave, and opposite it a small conical tube to be inserted in the ear.

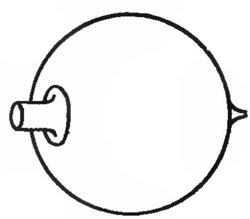


FIG. 70.

The free period of the enclosed mass of air determines the pitch of the tone to which the resonator will respond. To all other tones it remains practically silent. By means of a series of such resonators von Helmholtz was enabled to pick out the various overtones in the note of a piano string, and thereby to *analyze* a sound into its constituent tones.

133. Stationary Vibrations. Suppose one end of a long flexible cord be fixed and the other end be moved quickly up and down by the hand in a vertical plane. For each up and down motion of the hand a single pulse will run the length of the cord, be reflected at the fixed end and retrace the length of the cord, to be reflected again at the hand. In each case the reflection of the pulse in the cord will involve a *change of sign* both in the motion and in the nature of the disturbance itself, since a depression in the cord is returned as an elevation, and an elevation as a depression. After two reflections, therefore, the disturbance will have traversed the length of the cord twice and will be identical both in *direction* and *kind* with the original disturbance, and may be regarded as starting out anew.

If now the hand be maintained in simple harmonic motion, a series of harmonic waves will run along the cord and be reflected at the fixed end as before. Each wave after two reflections will coincide both in *direction* and *kind* with the new outgoing wave, if the time required to travel twice the length of the cord be some *whole* number, k, times the period of the motion maintained by the hand. In other words, if

$$\frac{2l}{V} = kT \tag{207}$$

where l is the length of the cord, V the velocity of the pulse along the cord, T the period of the motion, and k is *any integer* as 1, 2, 3, 4, etc.

Under the above condition it is clear that any disturbance however small will soon be increased sufficiently to set the cord into vibrations of wide amplitude at all points where the direct and reflected waves *coincide in phase.* Such points are called *antinodes.* At certain other points, however, the incoming and

outgoing waves meet in *opposite phase* and produce points of *minimum motion;* such points are called *nodes.* As a result of the superposition of the direct and reflected waves the cord is broken up into a series of vibrating loops, or ventral segments (Fig. 71), separated by points of minimum motion *N.*

Such a vibratory motion is called a *stationary vibration* or a *stationary wave.* The distance from one node to the next, or from one antinode to the next, is *one half wave length of the pulse in the cord.* Stationary vibrations may thus be set up in any medium capable of transmitting wave motion, and the phenomena of nodes and antinodes developed according to the principles just laid down. In all cases, the distance from *node to node*, or from *antinode to antinode*, is $\lambda/2$, *for the medium in question.* From *node* to *antinode* is $\lambda/4$.

Fig. 71.

In all cases of sustained tones, as those from organ pipes, tuning forks, piano, violin or guitar strings, the vibrating medium, whether air column, bar or string, is executing *stationary vibrations*, and consequently presents the characteristic feature of stationary waves, *i.e.*, *nodes* and *antinodes.*

134. Laws of Transverse Vibrations of Strings. A string fastened at the ends and vibrated transversely executes stationary vibrations as described in the previous article. The vibration of the string gives rise to a note of definite pitch, dependent upon the physical constants of the string and upon its mode of vibration. The condition for stationary vibration is

$$\frac{2\,l}{V} = kT = \frac{k}{n} \tag{208}$$

where n is the frequency of the note produced by the string. Of the various modes of vibration dependent upon the value of k, the simplest is that in which the string vibrates as a whole, with a node at each end. In this case $k = 1$, and

$$n = \frac{V}{2\,l} \tag{209}$$

This vibration is called the *fundamental vibration*, and the tone the *fundamental* or lowest tone given by the string.

It may be shown mathematically that the velocity of a *transverse wave* in a thin, flexible string, of density d and radius r when stretched by a *force* of T dynes, is given by the expression

$$V = \sqrt{\frac{T}{\pi r^2 d}} \tag{210}$$

If we substitute this value for the velocity in the equation for the frequency, we have

$$n = \frac{1}{2l}\sqrt{\frac{T}{\pi r^2 d}} \tag{211}$$

Hence the frequency of the fundamental tone emitted by a string vibrating transversely, varies

(*a*) Inversely as the length of the string.

(*b*) Inversely as the radius of the string.

(*c*) Inversely as the square root of the density of the string.

(*d*) Directly as the square root of the stretching force.

These laws are deduced theoretically from the case of a long, thin and perfectly flexible string, and are very nearly realized in the case of silk or gut strings. If metallic strings are used, the rigidity of the string acts as an added stretching force, thus making the frequency higher than the formula would indicate.

135. Melde's Experiment. If a thin, flexible cord some four meters long be attached to some convenient source of simple harmonic motion, and the stretching force be properly adjusted, the laws of vibrating strings may be verified by experiment.

In Melde's experiment one end of the cord is attached to one prong of a vertical tuning fork, and the other end carrying a pan is passed over a light pulley. If the prongs of the fork stand normally to the cord, and vibrate in the direction of its length, then the end of the cord will be displaced longitudinally at each vibration of the fork. At the forward swing the cord relaxes and drops down, rises and stretches tight on the backward swing of the fork, passes the position of rest and rises above on the next forward swing, falls to the middle on the

backward swing and to its lowest position again on the forward swing. The fork has thus made *two complete* vibrations, for a *single vibration of the cord*, or the cord in this position, *vibrates an octave lower than the fork.*

If the stretching force be properly adjusted by weights placed in the pan, the cord opens out into a wide spindle which remains fixed while the vibration continues. If the stretching force be reduced to one fourth its value, the velocity of the pulse will be one half its previous value, and the cord will now present two spindles with a node in the center. One ninth the original stretching force will give three spindles and two nodes. This verifies the law of the stretching forces.

If the fork be rotated about a vertical axis so that its vibrations are normal to the length of the cord, the cord will vibrate in unison with the fork, and for the original stretching force, it will divide into *two spindles*, where it previously vibrated in *one.* Since *each half* now vibrates in unison with the fork, while before the whole cord vibrated an octave *below* the fork, the law of lengths is demonstrated.

136. Segmental Vibration. It has been shown that a string may execute stationary vibrations under the condition that

$$\frac{2l}{V} = kT = \frac{k}{n}$$

This means that the frequency of the tone emitted by a string may depend upon *its mode of vibration*, as well as upon its *length*, *diameter*, *density*, or the stretching force to which it is subjected. In the case of a string sounding its fundamental, k is unity and the string *vibrates as a whole*, with nodes at the two ends. This is the simplest mode of vibration. The next simplest is when the string is divided into two segments with a node in the middle. In this case $k = 2$, and the frequency is double that of the fundamental. This tone may be drawn from a string by holding it at its middle point with a thin shaving of cork and bowing it lightly at about one ninth its length from one end. Here the string may be considered as made up of two strings of equal

length vibrating in unison, and producing a tone an octave above the fundamental. In a similar way the string may be broken up into three, four, five, six, or any number of equal segments, corresponding to the integral values of k. The frequency of the tone in each case is inversely as the length of the segments, and consequently if the string vibrate in three segments, the frequency is *three times* that of *the fundamental;* if it vibrate in fourths, the frequency is *four times* the fundamental; if in fifths, five times, etc.

The experimental demonstration of segmental vibration is most readily accomplished by means of a piece of piano wire about four meters long, tightly stretched between two bridges clamped to the top of a long table. One end should be attached to a key or screw in order to vary the stretching force at will. A series of stiff paper markers should indicate the aliquot parts of the string, as the thirds, fourths, sixths, etc. A thin shaving of cork should be slit and fixed upon the wire so as to slide freely upon it. Little riders of white and colored paper may be distributed along the wire, with the white ones at the aliquot points of division.

If the cork be held at one eighth the length of the wire and the bow be applied, gently at first, then more vigorously, the riders at the eighths will remain seated while all the others will be thrown off. The frequency of the tone emitted by the string will be eight times that of the fundamental. In a similar way the string may be made to vibrate in sixths, fifths, fourths, thirds and halves, the position of the cork marking a node in each case. If the cork be *not* upon some aliquot point, no satisfactory vibration and no note of definite pitch can be produced, hence *we conclude that a uniform string may vibrate as a whole, or in any number of equal parts, and the frequency of the note emitted will be proportional to the number of parts.*

137. Overtones. If the piano wire of the preceding experiment be vigorously bowed and then damped at one fourth its length, the note will be observed to change its character. The fundamental will disappear and the second octave, a note whose frequency is four times that of the fundamental, will be heard

instead. This shows that while the string was *vibrating as a whole*, it was also *vibrating in fourths*, and further, that the note emitted was made up of the *fundamental tone and the second octave*. By successively bowing the string, and damping it at the middle, and at one third its length, the first octave and the twelfth are found to be present when the string vibrates freely.

It thus appears that a string may at the same time vibrate as a whole and divide into several sets of equal segments, thus giving rise to the fundamental and also to tones whose frequencies are much higher than that of the fundamental. The higher tones thus obtained are termed *overtones*, or *upper partials*. In case their frequencies are exact multiples of that of the fundamental, the entire series are called *harmonics*, in which the fundamental is properly termed the *first*, the first octave the *second*, the twelfth the *third harmonic*, and so on, since these tones represent a series in which the frequencies are as 1 : 2 : 3 : 4 : 5, etc.

If we consider the key of C, the first ten harmonics, counting the fundamental as the first, are C (1), c (2), g (3), c_1 (4), e_1 (5), g_1 (6), c_2 (8), d_2 (9), e_2 (10); where the *seventh harmonic* lies between a_1 and b_1 and may be represented by b_1 flat.

Of these tones it will be observed that the first six harmonics are consonant tones, and if sounded together would produce the effect of a *perfect major chord*. In a string of uniform dimensions and homogeneous structure, the frequencies of the *upper partials* approach very nearly the exact relation demanded for harmonics. It is for this reason that the music from stringed instruments is so rich and pleasing.

Again, it is clear that if the string be bowed or struck at its *middle point*, that point cannot by any possibility be a node. Hence all partial tones which require the presence of a node at the middle of the string must of necessity be absent. The discordant effect of the seventh and ninth harmonics of a string are avoided in the case of a piano by having the hammer strike the wire at a distance of a little less than one seventh the length of the string from one end.

Problems

1. Show that the right and left hand members of the equation

$$V = \sqrt{\frac{e}{d}}$$

are of the same dimensions.

2. Assuming the velocity of sound in air at 0° C to be 331.36 m per second, calculate the velocity in hydrogen at the same temperature, having given the mass of one liter of hydrogen = 0.0896 g. *Ans.* 1259 m per sec.

3. Find the temperature at which the velocity of sound in air is 356 m per second. *Ans.* 40.°6 C.

4. The flash of a hunter's gun is seen and after 5 sec the sound is heard. Required the distance from the observer to the hunter, the temperature being 22° C. *Ans.* 1723.46 m.

5. Colladon and Sturm measured the velocity of sound in the waters of Lake Geneva, and found that it traveled 1435 m per second, the temperature being 8°.1 C. Compute the coefficient of elasticity for water at this temperature. *Ans.* 20.59×10^9 dynes per cm².

6. A wire 50 cm in length and of mass 80 g is stretched so that it makes 100 complete vibrations per second. Compute the stretching force. *Ans.* 16×10^7 dynes.

7. A string is attached to one prong of a tuning fork and after passing over a smooth peg is stretched with a force of 32×10^3 dynes. When the string is parallel to the motion of the fork it vibrates steadily in three segments. What stretching force is required to make it divide into two segments? Into five segments? *Ans.* (*a*) 72×10^3 dynes.
(*b*) 11.52×10^3 dynes.

8. What stretching force is needed to have the above-mentioned string divide into eight segments, when the string stands at right angles to the motion of the fork? *Ans.* (*a*) 18×10^3 dynes.

9. Determine the vibration frequency of an air particle in a sound wave 10 m long ($t = 20°$ C). *Ans.* $n = 34.348$ per sec.

10. If the first syllable of an echo reaches the ear 3 sec after the spoken word, how far distant is the reflecting surface? ($t = 20°$ C.) *Ans.* 515.22 m.

11. A stone is dropped into a well and the sound of the splash is heard after 5 sec. How deep is the well, if the temperature be 10° C? *Ans.* 107.5 m.

12. An open organ pipe 120 cm in length is tuned correctly when the room temperature is 20° C. What will be the change in its frequency when the temperature rises to 32° C? *Ans.* 3 vibrations per sec.

13. A workman strikes a blow with a hammer upon one end of an empty iron water pipe, 600 m long. A second workman placing one ear against the other end hears two sounds. How far apart are they in time? Temperature 20° C. $\left(V \text{ for sound in cast iron} = 5000 \frac{\text{m}}{\text{sec}}.\right)$ *Ans.* 1.62 sec apart.

14. A horizontal string carrying a small globule of mercury is viewed through a lens fastened to one prong of a tuning fork, placed at right angles to the string and vibrating horizontally. The fork has a frequency of 128. The fork is bowed and the string is tuned until the ellipse seen through the lens makes one complete rotation in 6 sec. If the stretching force be increased, the Lissajous' figure rotates faster. What is the per cent of error if this tuning be assumed as correct? *Ans.* 0.13 %

CHAPTER XVI

VIBRATION OF AIR IN PIPES AND CAVITIES

138. Vibration of Air Columns. In many musical instruments the vibrating body is a column of air in a pipe. Although the shape of the column and the mode of excitation may vary, yet the general principles of vibrating bodies will apply with but slight modification. When a series of similar pipes of the same diameter but of different lengths are sounded by blowing in turn across the ends of each, it will be found that the frequencies of the sounds produced are practically inversely as the length; that is, a slender pipe 10 cm long will give a note approximately one octave higher than a similar pipe 20 cm long, and two octaves higher than one 40 cm long.

If a tuning fork be held over a vertical pipe, the lower end of which is connected with a water supply for varying the length of the enclosed air column, it will be found that for a certain level of the water the air column in the pipe responds loudly to the vibrations of the fork. If pipes of different diameters are used, it will be found that under similar conditions the length of pipe responding to a given fork is nearly constant, diminishing slightly as the diameter increases. Again, if a closed pipe 20 cm long respond to a given fork, it will be found that an open pipe of the same diameter and same length will respond to a fork an octave higher than the first fork. This shows that *the pitch of an open pipe is an octave higher than that of a closed pipe of the same length.*

139. Length of Organ Pipe and Wave Length of Fundamental Tone. (*a*) *Open pipe.* Suppose an open pipe (Fig. 72) have placed before one end a tuning fork or other suitable vibrator, which sends a series of sound waves into the pipe. Then *for*

the pipe and the tuning fork to vibrate in unison, it is necessary for the reflected wave to return to the end B, in the proper phase to unite with the outgoing wave. This means that when the fork starts to swing from a'' to a' a condensation is sent into the pipe and runs the length BA while the prong moves the distance $a''a'$. At the open end A the condensation is reflected as a rarefaction (Art. 118), which starts into the tube at A at the same instant that a rarefaction enters at B, due to the backward motion of the prong from a'. These two rarefactions meet at the center of the tube, producing a *double rarefaction* for an instant, and pass on to the open ends to be reflected as condensations, at the instant the fork begins a second swing from left to right. The condensation at B unites with a new condensation from the fork, and the combined condensations run in at B while the reflected condensation enters at A. The two condensations meet at the middle, forming a *double condensation* at that point, just half a period later than the double rarefaction.

FIG. 72.

The disturbance is thus seen to travel the length of the pipe twice during one complete vibration of the fork, or, for the fundamental tone of the open pipe,

$$\frac{2\,l}{V} = T \tag{212}$$

whence $$2\,l = VT = \lambda$$

This shows that for an open organ pipe *the wave length of the fundamental is twice the length of the pipe.*

(*b*) *Closed pipe.* In the closed pipe the fork on its swing from left to right sends in a condensation which runs to the closed end and being reflected as a condensation runs back to the open end, where it emerges and combines with the outgoing condensation caused by the fork on its swing from right to left. At the same time the emerging condensation is reflected *into* the pipe at the *open end* as a rarefaction which combines with the rarefaction left in the rear of the fork on its

passage from right to left. The condensation or the rarefaction has in each case run the length of the pipe *twice* during *half* a vibration of the fork, or for unison with the fundamental,

$$\frac{4\,l}{V} = T \tag{213}$$

whence $$4\,l = VT = \lambda$$

That is, the wave length of the fundamental of a closed organ pipe is *four times the length of the pipe.* If we compare this result with that obtained for the open pipe, we see that the wave length of the closed pipe is double that of an open pipe of the same length, or the fundamental of a closed pipe is an octave lower than that of an open pipe of the same length.

140. Nodes in Open and Closed Organ Pipes. In the open pipe it was shown that a node existed in the middle, at which point there existed alternately double rarefactions and double condensations at intervals of half a period. A node therefore in a vibrating air column is to be considered *as a place of maximum change of density, but of minimum motion.* In an open pipe there is *always an antinode at each end,* since at the open end the motion is unrestricted. For the fundamental in an open pipe, therefore, there is a node in the middle and an antinode at each end, or the pipe contains *one half wave length* (Fig. 73 *A*).

If the pipe be blown more strongly, it gives its first octave, the air column breaks up into segments, with a node at one quarter the length of the pipe from each end, and an antinode at the middle and at each end. In this case the pipe contains *two half wave lengths,* and the corresponding note is an octave above the fundamental (Fig. 73 *B*).

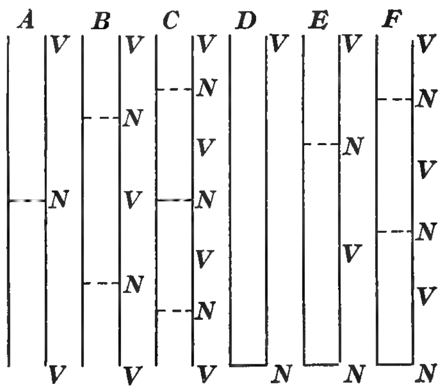


FIG. 73.

The second overtone is given by the air column forming three nodes, one at *one sixth* the length of the pipe from either end and at the middle, with antinodes between. In this case the pipe contains *three half wave lengths*, and the frequency of the tone is three times that of the fundamental (Fig. 73 *C*), and so on for higher tones. *Hence in an open pipe all the overtones are present.*

In a closed pipe, *there is always a node at the closed end*, since the air is at rest there, and as usual there is an antinode at the open end. For the fundamental the pipe contains *one fourth wave length* (Fig. 73 *D*).

The first overtone in the closed pipe is given by the air column forming a new node, at one third the distance from the open end. The pipe thus contains *three fourth wave lengths*, and the tone has three times the frequency of the fundamental (Fig. 73 *E*). The second overtone produces a node at *one fifth* and *three fifths* the length of the pipe from the open end. The pipe contains *five fourth wave lengths*, and the frequency of the tone is *five* times that of the fundamental (Fig. 73 *F*). *Hence in closed pipes only those overtones are present whose vibration frequencies correspond to the odd multiples of the fundamental.*

141. Quality of Sound. By means of his analysis of musical sounds von Helmholtz decided that the *quality* of a sound depends upon the number of overtones associated with the fundamental and upon their relative intensities, and is independent of their differences in phase. Quality of sound depends upon the form of the sound wave. In the ear the various constituents of a complex wave are separated and noted, and the effects of the various combinations distinguished. Von Helmholtz showed not only by direct analysis, but also by synthesis, that the sounds of certain musical instruments consist of definite overtones combined with the fundamental. By means of a series of tuning forks each of which gave a simple tone, he was able successfully to reproduce the notes of various musical instruments, and even to imitate most of the vowel sounds of the human voice.

An admirable instrument for the analysis of sound is found in the manometric capsule devised by Koenig. A cylindrical box (Fig. 74) is divided into separate, air-tight compartments by a flexible diaphragm, *D*, of thin rubber, or goldbeater's skin. Into one of the compartments, *A*, are introduced the sound waves by means of the funnel *M*. The compartment *B* contains illuminating gas, which enters through the tube *C*, and is ignited at the tip. If a condensation impinge upon the diaphragm, the gas in *B* is compressed and the flame leaps up; if a rarefaction enter *A*, the pressure in *B* is less and the flame is drawn down. If a musical note be sung into the funnel, the flame vibrates in unison with the air particles in *A*, and if it be viewed in a rotating mirror, the eye can determine at once the nature and constitution of the sound. In Fig. 75 the upper picture represents the appearance in the mirror when a simple tone is sounded in the funnel. The middle line represents the appearance of the flame when the octave of the first note is sounded, and the third shows the effect of combining the two. Manometric capsules may be attached to each of a series of resonators and the combination affords a means of instantly determining the composition of any note sounded in its vicinity.

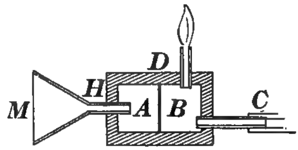


FIG. 74.

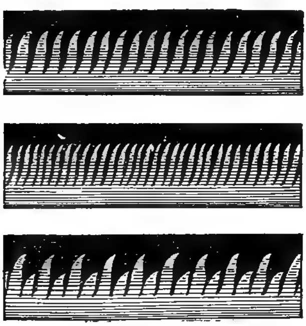


FIG. 75.

142. Kundt's Experiment.[1] The principle of resonance has been ingeniously applied by Kundt to the measurement of the velocity of sound in solids or in gases. A long glass tube (Fig. 76), some 6 cm in diameter, is furnished at one end with a loosely fitting piston, and has the other end closed by a sheet

[1] *For experimental details of Kundt's Experiment, see Manual, Exercise 34.*

of thin rubber. A rod of brass or other metal is held by its middle point in a vise and one end is furnished with a disk of stiff paper which rests against the rubber membrane of the

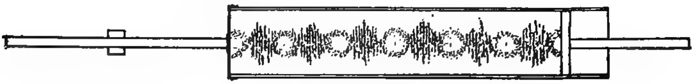


FIG. 76.

glass tube. The inside of the tube is dusted with fine cork filings or amorphous silica.

When the rod is stroked with a piece of chamois skin covered with a little rosin, it gives a loud clear note. It vibrates longitudinally, with a node in the center, after the manner of the air in an open organ pipe. The paper disk communicates the vibrations of the rod to the air in the tube, and when the length of the enclosed air column has been properly adjusted by means of the piston, the cork dust is tossed about into little heaps owing to the resonance of the air with the note emitted by the rod.

As we have already seen, the conditions of resonance demand that the path run over by the disturbance from the end of the rod to the end of the air column and back again, must be some even number of wave lengths, since the tube is closed at each end, and the motion is twice reflected with change of sign in direction, but without change of sign in condensation. When resonance has been established, the powder in the tube assumes the appearance shown in Fig. 76. The nodes, being points of minimum motion, are marked by small circles of the cork filings, while the antinodes are shown by transverse heaps of dust, where it has fallen at the cessation of the sound.

In Art. 133 it was shown that the distance from *node* to *node* is a half wave length of the disturbance in the medium. The nodal lengths in the air may be readily measured from the circles of powder, and the half wave length of the sound in air computed. Since the rod behaves as an open organ pipe sounding its fundamental, it follows that the length of the rod is *one half wave length of the sound in brass.* If V_a and V_b, λ_a and λ_b

represent the velocities and wave lengths of sound in air and brass respectively, then, since the period is the same in each case, we have

$$\frac{V_b}{V_a} = \frac{\frac{\lambda_b}{2}}{\frac{\lambda_a}{2}} \tag{214}$$

or

$$V_b = V_a \frac{\lambda_b}{\lambda_a} \tag{215}$$

For gases, an additional tube with powder is fitted up and placed in contact with the other end of the rod. From the measured lengths of the nodal distances in the gas and in air the computation is made as given above.

***143. Mouthpieces.** The various forms of wind instruments differ chiefly in the mode of excitation of vibration of the enclosed column of air. That part of the instrument in which such vibration is excited is called the mouthpiece. Mouthpieces may be divided into three classes.

(*a*) Those in which the air is blown *across* a sharp edge or across an opening, as in the common tin whistle, the organ pipe or the flute.

(*b*) Those in which the air is forced *through* an opening, either partially closed by an elastic tongue or reed which swings through, as in the common cabinet organ, harmonica, accordion, etc., or closed by a reed which shuts *down upon* the opening as in the clarionet, oboe and bassoon.

(*c*) Those in which the air is forced through a slit formed of two elastic membranes. This form of mouthpiece is made by cutting off the ends of a wooden tube obliquely on opposite sides and tying two strips of thin rubber over the faces so formed, so as to leave a narrow slit along their line of junction. If air be blown through the slit, a note will be produced, whose pitch will be modified by the body of air in the tube.

***144. Vocal Organs.** Of all musical instruments the larynx, the organ of human speech and song, is the most wonderful,

both on account of its simplicity as well as for its extreme delicacy and range.

The larynx may be briefly described as a box formed by three plates of articulated cartilage which are moved by muscles, placed at the upper end of the trachea or windpipe.

The base of the larynx is formed of a large ring of cartilage called the *cricoid* (ring shaped) cartilage. Attached to this is the *thyroid* or shield-shaped cartilage, which is bent in the shape of a *V*, and fastened to the edges of the *cricoid* ring by its sides, the point of the *V* being turned to the front, forming the projection on the front of the throat known as "Adam's apple." At the back of the *cricoid* are fastened two small pointed cartilages, the *arytenoid* (funnel shaped) cartilages. Stretching between the *arytenoid* cartilages to the inner sides of the *V*-shaped *thyroid* are two elastic membranes, one fastened to each leg of the *V*. These are the *vocal chords*, *cc*

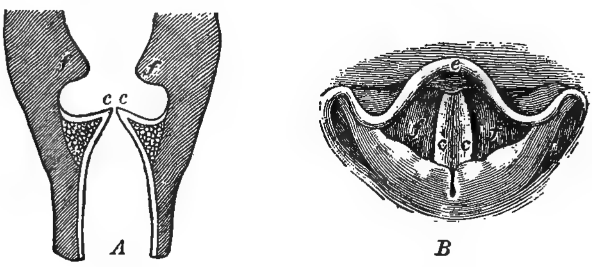


FIG. 77.

(Fig. 77 *B*). Just above these are two folds of mucous membrane, known as the *false vocal chords*, *ff* (Fig. 77 *A*).

When the peak of the *thyroid* is not drawn down, the vocal chords are lax and the breath passes freely between them as in ordinary breathing. But when the peak of the thyroid is pulled down by the muscles attached to it, the vocal chords are stretched, the arytenoid cartilages move nearer to each other, and the thin, sharp edges of the vocal chords form a narrow slit across the windpipe, through which the air is forced, causing them to vibrate as the rubber membranes in the third form of mouthpiece described in Art. 143.

The pitch of the tone produced by the vocal chords depends upon their size, length and tension. The tension is controlled by the attached muscles, so that the singer can vary at will the pitch of the tone produced.

The quality of the sound produced is modified by the resonant qualities of the cavities of the mouth, throat and nasal passages.

*145. **The Ear.** The human ear consists of three well-marked divisions, termed the *external*, *middle* and *internal ear*. The

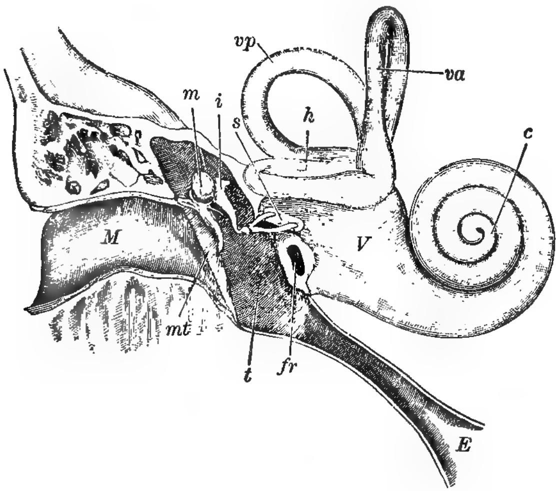


Fig. 78.

external part includes the familar appendage at the side of the head, and the auditory canal, or *meatus*, *M* (Fig. 78), in which is shown a section through the right ear.

The meatus is closed by the *tympanum*, or eardrum, *mt*. This separates the external from the middle ear, and acts as a receiving membrane against which the sound waves impinge.

In the middle ear is found a chain of three small bones: the hammer (*malleus*), *m*; the anvil (*incus*), *i*; and the stirrup (*stapes*), *s*.

The handle of the hammer is attached to the tympanum and

the base of the stirrup rests against the membrane of the *oval window*, a small oval opening into the inner ear. The middle ear communicates with the external air by means of the *Eustachian tube E*, which leads to the upper part of the throat.

The internal ear, or labyrinth, is seated deep in the skull, and communicates with the middle ear by two openings in its bony case, the *oval window*, and the *round window, fr.* It consists of three parts: the *vestibule, V;* the *semicircular canals, h, vp, va;* and the *cochlea*, or snail shell, *c*, all of which are filled with a watery fluid.

The cochlea is a tapering tube, coiled up like a snail shell, and divided longitudinally into two compartments. These are formed by a bony partition, extending out from the axis of the spiral, and two membranes joined to its edge and attached to the walls of the tube. In the fluid of the cochlea between the two membranes are found about 3000 rods of different length, known as the rods of Corti. One of the membranes called the *basilar membrane*, consists of from 18,000 to 24,000 fibers, radially stretched strings, varying in length from the top to the bottom of the cochlea. A nerve filament from the brain is supposed to be connected to each of these fibers. A large number of stiff, elastic hairs are also found floating in the fluid of the vestibule and attached to the membrane on its sides.

The process of hearing begins in the transmission of the sound waves to the drum of the ear through the air. From the eardrum the vibratory motion is transferred by the small bones, to the yielding membrane of the oval window, and transmitted to the enclosed liquid of the vestibule and the cochlea, by which it is transmitted to the fibers of the basilar membrane. It is thought that the function of these fibers is to resolve sounds into their components and to report the individual components to the brain. It is assumed that such resolution is effected through resonance, each fiber responding to some specific tone.

The semicircular canals *h*, *va*, and *vp* (Fig. 78) are so placed that one lies in a horizontal plane, the second in a vertical plane from front to back, and the third in a vertical plane from right

to left. These canals seem to have no connection with the sense of sound, but are regarded as the organ of the sense of equilibrium. Animals from which one or more of these tubes have been removed seem to experience great difficulty in maintaining their equilibrium.

HEAT

INTRODUCTION

CHAPTER XVII

NATURE OF HEAT

146. Nature of Heat. When a vessel containing cold water is placed over a fire, the water first becomes warm, then hot, and finally begins to boil, owing to the passage of heat from the fire to the water. After the water has begun to boil, it will not become appreciably hotter, no matter how long the boiling may be continued. If the fire be made more intense, the water will only boil more rapidly. There is no reason to suppose that heat ceases to pass into the water when the boiling begins. On the contrary, it seems clear that large quantities of heat are absorbed in order to keep the water boiling, and that the only effect which this heat produces is the continuous formation of steam from the boiling water.

If the experiment be begun with a vessel filled with pieces of ice, the ice melts, but if heat be added slowly and the contents of the vessel be continually stirred, the mixture of ice and water remains practically as cold as the ice, so long as there is any ice present. Only after the ice is all melted does the water begin to grow sensibly warmer.

From this we see that the addition of what we call heat to a body produces two distinct effects: (*a*) A change in the condition of the body affecting the sensation of warmth and cold. Such expressions as "cold," "cool," "warm," "tepid," and "hot," all refer primarily to this sensation, or they characterize the degree of hotness of the body. In scientific language the thermal condition of a body is designated as its *temperature*, and the hotter a body is, the higher its temperature is said to be.

(*b*) A change of state, such as the formation of water from ice or steam from water.

The addition of heat to a body produces many other effects. For example, if a thermometer be placed in hot water, the mercury rises in the tube, showing that the volume of the mercury increases on the addition of heat.

Again, if instead of the thermometer we place in the hot water a thin glass flask completely filled with air or any other fluid, and closed by a stopper to prevent change of volume, we shall see that one of two things will happen: either the stopper will be driven out or the flask will be broken. This shows that the pressure of an enclosed mass of fluid increases on heating, provided the volume be kept constant.

In fact, almost every physical property of a body will change when the body is heated, but the examples cited are sufficient to show that change of temperature is only one of the many effects which heat can produce, and that heat and temperature are two entirely different physical quantities.

Heat was formerly supposed to be a subtle, imponderable fluid called "caloric," which upon entering a body produced the effects considered above (Black,[1] 1728–1799). Many of the terms used in the following pages date back to the caloric theory. At the close of the eighteenth century, however, a number of experiments by Count Rumford[2] (1753–1814) and Sir Humphry Davy[3] (1778–1829) proved definitely that *heat may be produced by work done against friction*, and that *the quantity of heat produced bears a constant ratio to the quantity of work expended in overcoming friction.*

These experiments disposed of the caloric theory of heat once for all, and by connecting the production of heat directly with the known laws of mechanics, placed this branch of physics upon a sound scientific basis. At present *heat is known to be a form of energy which shows itself chiefly through the effects of temperature changes produced by it in different bodies.*

[1] Black, *Lectures on the Elements of Chemistry*, vol. i, p. 156; 1803.

[2] Rumford, *Phil. Trans.* 1798 and 1799.

[3] Davy, *Complete Works*, vol. ii, p. 11.

147. Molecular Theory of Heat. The view that heat is a form of energy is closely connected with the molecular theory of the constitution of matter. In this theory it is assumed that the molecules of a body are in continuous and irregular motion. In a solid this molecular motion is restricted to a small space about the position of equilibrium of the molecule which, in general, may be considered as fixed in the body. In liquids there is more molecular freedom, so that the molecules are able to slip past each other and move about from point to point within the liquid, as is shown in the phenomena of diffusion (Art. 98). In a gas the molecules are assumed to have great freedom of motion and to be held together only by the application of some external force. In gases of ordinary density the mutual force of attraction between the individual molecules is conceived to be extremely small, and to exert little or no tendency to draw the mass of gas together, until the gas has been compressed by external force almost to the point of liquefaction.

It is also to be noted that the motion of the molecules is assumed to be not only *continuous* and *irregular*, but *blind* and *undirected* as well. By this we mean that the motion of the molecules of an enclosed gas is not comparable to that of persons in a crowd or of bees in a swarm, where the motion, while both continuous and irregular, is yet such that the individual avoids collision with his fellows. The molecules of the gas, however, are continually colliding with the walls of the enclosing vessel and also with each other.

The average distance traversed by a molecule between collisions is termed *the mean free path of the molecule.* Obviously this mean free path decreases as the density of the gas increases; when the gas has been compressed until liquefaction occurs the mean free path is greatly restricted, while in the case of a solid the molecule is closely hemmed in by its neighbors, and its motion becomes in general an irregular vibration about its position of equilibrium, in which it is held by the molecular attraction of its fellows. However, diffusion has been observed between solids, especially at higher temperatures, showing that the molecules of a solid have still a certain amount of freedom.

The energy of the molecule, like the energy of the whole mass, may take either the potential or the kinetic form. Similarly, the kinetic energy of the molecule may be either translational or rotational in character. If heat be added to a body, there results an increase in the molecular energy of the body at the point at which the heat enters the body. This increase of energy soon becomes distributed throughout the body, owing to the constant collisions of the molecules. If the average molecular kinetic energy per unit volume of the body has become the same in all parts of the body, the body is said to be uniformly heated. The rise in the temperature of a body is thus explained by an increase in the kinetic energy of the molecules.

If we strike a piece of iron with a heavy hammer, the mass kinetic energy of the hammer suddenly disappears and in its stead the molecules of the hammer and of the piece of iron are thrown into more violent agitation, or the hammer and the iron are heated. In this way we see that the mass kinetic energy of the hammer has been transformed into molecular kinetic energy in the hammer and in the iron.

Although the molecular theory of heat furnishes a satisfactory explanation for the phenomena of heat, yet the laws and relations derived in the following paragraphs are entirely independent of any theoretical explanation of the processes of heating and cooling.

TEMPERATURE

CHAPTER XVIII

THERMOMETRY

148. Temperature. After a body has either gained or lost heat, its thermal condition has undergone a certain change. Neglecting for the present any possible change of state, the most noticeable change in the body produced by this gain or loss in heat is a change in its temperature. It therefore becomes a matter of first importance to arrive at some reliable means for the measurement of this temperature change.

Our temperature sense furnishes us an approximate idea of the thermal condition of a body, although this estimate is greatly influenced by other factors. Thus on a cold day a piece of metal feels very cold to the touch, while a piece of wood seems less cold and a piece of wool may even appear warm. On a hot day these sensations would be exactly reversed, since the metal would then seem hot, the wood less so and the wool would appear cool. From this it is clear that our sensations of heat and cold depend not so much upon the actual temperature of the body in question, as upon the rate at which heat passes from the hand to the body, or from the body to the hand in either case. The apparent differences in temperature are due to the fact that this rate of transmission of heat from or to the hand in the foregoing examples is much greater in the case of metals than in the case of such substances as wood or wool.

A more reliable and accurate method of comparing temperatures is found in the measurement of some one of the many physical changes accompanying a change in the temperature of a body. *In this sense temperature is the physical measure, on an arbitrary scale, of that condition of a body, capable of affecting our sense of warmth and cold.* An instrument used for measuring temperature is called a thermometer, or in the case of very high temperatures, a pyrometer.

It is evident that a thermometer indicates directly, simply its own temperature, and only indirectly the temperature of the medium in which it is immersed. However, the *law of equalization of temperature* states that, when two bodies are in close thermal contact, and do not lose heat rapidly to their surroundings, nor gain heat from outside sources, any difference in temperature will soon disappear unless one of the bodies either gains or loses heat more rapidly than the other. Consequently a thermometer immersed in a fluid will soon assume the temperature of the fluid, and its readings, after becoming constant, may safely be assumed to indicate the temperature of the fluid in which it is immersed.

149. The Mercury-in-glass Thermometer. The most common form of thermometer is the mercury-in-glass thermometer, or a thermometer in which mercury is used as the thermometric substance, whose expansion in contrast to the expansion of the enclosing glass bulb is to be used to measure temperature changes. The *principle of temperature measurement* is, that changes in temperature, corresponding to equal changes in the apparent volume of mercury in glass, shall be called equal changes in temperature. As a thermometric substance, mercury has certain advantages over other liquids:

(*a*) It is easily obtained in a pure state.

(*b*) It does not stick to the glass.

(*c*) It is opaque and easily seen in a capillary glass tube.

(*d*) Its expansion with change of temperature is quite large.

(*e*) To each temperature there corresponds one and but one definite volume.

(*f*) It requires but a small amount of heat to raise its temperature through a given range, and so the introduction of the thermometer does not appreciably change the temperature of the substance whose temperature is to be measured.

In order to render small changes in volume noticeable, the mercury is enclosed in a glass bulb ending in a tube of capillary bore. The sensitiveness of the thermometer is greater, the larger the bulb and the finer the capillary tube.

150. Limitations of the Mercury-in-glass Thermometer. Mercury freezes at $-39°$ C, and consequently cannot be used to measure temperatures below this point. Instead of mercury, alcohol, toluene or pentane are used for low-temperature measurements. Pentane remains liquid even at $-200°$ C. Since these liquids expand at rates very different from that of mercury, a careful calibration of such thermometers is necessary. Mercury boils at 356°.7 C, under atmospheric pressure, but the boiling may be prevented by an increase of pressure. In order, therefore, to use mercury thermometers for still higher temperatures, it is only necessary to fill the space above the mercury in the tube of the thermometer with some inert gas, usually nitrogen. Upon the expansion of the mercury the inclosed nitrogen is compressed and the mercury is prevented from boiling. Mercury-in-glass thermometers reading up to 500° C are not uncommon, the only limit being the strength of the glass envelope. Glass becomes soft at temperatures higher than 550° C, but by using tubes of fused quartz which softens at about 1100° C, thermometers reading up to 700° C have been made.

Another limitation of the mercury-in-glass thermometer is due to the nature of the glass envelope itself. Glass bulbs when blown, contract slowly for months and even years. This slow contraction changes the volume of mercury in the bulb, and so changes all readings of the thermometer by a certain amount. In this way a thermometer comes to read too high, and correction for this error must be made. This defect may be largely overcome by allowing the bulbs to lie several years before filling them.

A second defect of the glass bulb arises from sudden and large temperature variations. Thus if a thermometer be placed first in ice water, then in boiling water, and again in ice water, the two readings in the ice water will not agree. This is due to the fact that the glass bulb, after having been expanded by the boiling water, does not return at once to its original volume when again immersed in the ice water. Since all readings are now too low, this defect is termed the "depression of the zero point." It may be almost entirely eliminated by the use of a special glass for making the bulbs.

151. Other Forms of Thermometer. As we shall see later, many other forms of thermometer have been devised and used for accurate scientific investigation or for use under conditions which would render the mercury thermometer unsuitable. The most satisfactory standard for temperature measurement is the change in pressure of a definite volume of hydrogen gas, under known conditions, and this thermometer, known as the standard hydrogen thermometer, has been generally accepted among scientific men (Art. 165).

Of the various other forms of thermometer used for special purposes, a few will be mentioned here, but each will be more fully described in connection with the physical principle involved in its construction. Among these are the strain thermometer, depending upon the linear expansion of solids; the gas thermometer, based upon the increase in volume or the increase in pressure of a mass of gas when heated; the resistance thermometer, involving the increase of electrical resistance with increase of temperature; the thermo-couple, depending upon the principles of thermo-electricity; and finally the radiation thermometer, based upon the laws of radiation, by means of which it has been rendered possible to measure the temperature of a remote source of energy, as, for example, the sun or the moon, whose radiations come to us over vast distances.

152. Centigrade and Fahrenheit Scales. After the thermometric substance and the principle of temperature measurement have been agreed upon, we are still free to select not only the value of a degree, or unit of change of temperature, but also the zero point from which to count.

It has been shown by experiment that the temperatures of pure melting ice, and of steam issuing from boiling water, are constant under a given pressure. These two temperatures, corresponding to a barometric pressure of 76 cm of mercury, are called the "freezing point" and the "boiling point." They are known as the fixed points of a thermometer, and the temperature difference between them is called the "fundamental interval."[1]

[1] *For experimental determination of the fixed points, see Manual, Exercise 37.*

In all scientific work the *Centigrade scale* is used, which divides the fundamental interval into 100 equal parts or degrees, and takes the freezing point as the zero of the scale. All temperatures below this are written with the negative sign. The subdivision into 100 parts was first suggested by Celsius of Upsala, in 1742. For this reason, this scale is sometimes called the "Celsius scale," although he chose the boiling point as zero and the freezing point as 100.

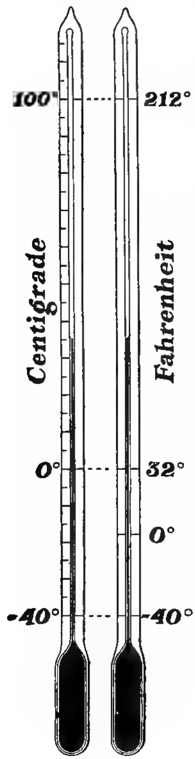


FIG. 79.

In English-speaking countries a different scale, invented by Fahrenheit of Danzig (1686–1736), is used, and it is called the *Fahrenheit scale.* In this the fundamental interval is divided into 180°, and the normal freezing point called 32°. It follows that the temperature of steam over water boiling under normal atmospheric pressure is 212° F.

The reading for a given temperature on one of these scales may easily be changed to the corresponding reading on the other scale. Call F the reading on the Fahrenheit, and C the reading on the Centigrade scale; then, since there are $F-32$ and C scale parts above or below the freezing point, we have (Fig. 79)

$$\frac{F-32}{C}=\frac{180}{100} \tag{216}$$

or

$$F=\tfrac{9}{5}C+32 \text{ and}$$
$$C=\tfrac{5}{9}(F-32) \tag{217}$$

In a third scale, called the *Réaumur scale*, the fundamental interval is divided into 80 parts, with the freezing point as zero. This scale, however, is little used in this country.

***153. Maximum and Mimimum Thermometers.** Frequently it is important to know only the highest or lowest temperature reached during a certain time interval. In Six's thermometer (Fig. 80) the large bulb A is filled with glycerine, while the

narrower tube is bent in a U-shape and partly filled with mercury. On the other end of the U-tube there is a small bulb *B*, partially evacuated, and also containing some glycerine. In the glycerine above the mercury, on both sides of the U-tube, two thin, short iron wires are placed, each carrying a small spring which presses against the wall of the tube. With rising temperature the mercury on the left-hand side pushes the wire above it upward, but on receding, as the temperature falls again, leaves it in the highest position reached. With decreasing temperature the mercury on the right-hand side rises and pushes its wire up until the minimum temperature is reached. The lower ends of the wires on the left and right hand sides, respectively, then give the maximum and minimum temperatures reached since the last setting. A new setting is made by pulling the wires down to the surface of the mercury, in each case by means of a magnet.

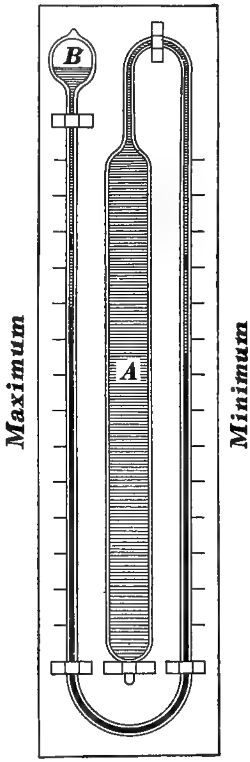


Fig. 80.

The United States Weather Service uses two separate thermometers. In the maximum thermometer the bore of the tube is constricted near the bulb. With rising temperature the mercury pushes through the constriction, but as soon as the temperature falls, the mercury breaks at this point, leaving the whole thread in the tube. Its reading gives directly the maximum temperature. The minimum thermometer is an alcohol thermometer, carrying within the alcohol a thin iron wire, which, by the surface tension of the alcohol, is pulled back to the lowest point reached. The alcohol on rising flows past the wire, leaving it at its lowest position. The end of the rod nearest the alcohol surface gives the minimum reading. Both thermometers are placed nearly horizontal (Fig. 81). The

maximum thermometer is reset by whirling it around the pin at its upper end. The centrifugal force drives the mercury back into the bulb. To replace the index in the minimum thermometer next to the alcohol surface, the bulb of the thermometer is raised until the wire slides forward.

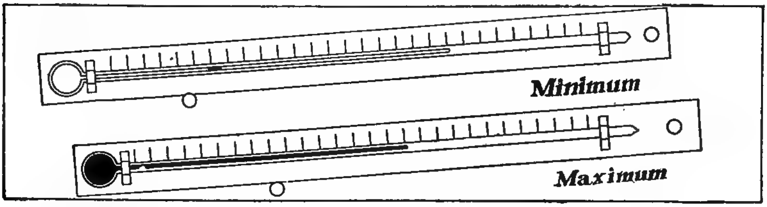


FIG. 81.

Clinical thermometers are also maximum reading thermometers, in which the mercury, broken by the constricted tube, gives the maximum temperature indicated by the instrument while in contact with the body of the patient. The mercury is brought back into the bulb by centrifugal force.

CHAPTER XIX

EXPANSION

154. Linear Expansion of Solids. In solids changes in length due to a change in temperature are, in general, very small. This change, however, may be shown, greatly magnified, by clamping one end of a metal bar (Fig. 82) about one meter long to a solid support and allowing the other end to rest upon a

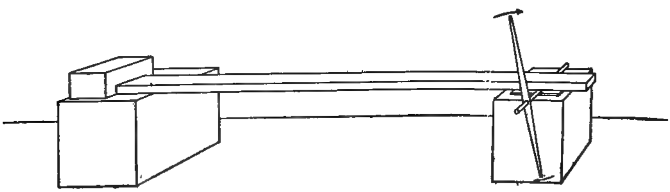


FIG. 82.

fine cambric needle, placed at right angles to the bar and free to roll upon a piece of plate glass. A long light pointer of paper attached to the needle will show its motion very clearly. On heating the bar its free end will advance slightly, rolling the needle forward, and its motion, greatly magnified, is shown by the paper pointer. When allowed to cool, the bar contracts, and the pointer moves in the opposite direction.

For the accurate determination of the linear expansion of solids, very refined methods are required. The best of these are optical methods involving interference of light, and the change in length caused by a change in temperature is expressed in terms of the wave length of the light used.[1]

If L_1 and L_2 represent the lengths of a solid bar, at temperatures t_1 and t_2, *then the change in length*, $L_2 - L_1$, is proportional to the original length L_1 and is nearly proportional to the

[1] *For a simple method for measuring linear expansion, see Manual, Exercise 38.*

change in temperature $t_2 - t_1$. This is expressed by the equation

$$L_2 - L_1 = \alpha L_1 (t_2 - t_1) \tag{218}$$

or

$$\alpha = \frac{L_2 - L_1}{L_1 (t_2 - t_1)} \tag{219}$$

The proportionality factor α is called *the coefficient of linear expansion*, and may be defined as the *change in length per unit length per degree*. The length at 0° C is usually taken as the original length, so that the length L_t, at any other temperature $t°$, is given by the equation

$$L_t = L_0 (1 + \alpha t) \tag{220}$$

However, α is not quite constant, but in general increases with increase of temperature, so that at higher temperatures the expansion per degree is somewhat larger than at lower temperatures. In any determination of this physical quantity, therefore, it should always be stated between what limits the observations were made.

TABLE VI

MEAN COEFFICIENT OF LINEAR EXPANSION BETWEEN 0° AND 100° C PER DEGREE C

Substance	$\alpha \times 10^6$	Substance	$\alpha \times 10^6$
Copper	17.1	Glass	7 to 9
Iron	12.1	Jena thermometer glass .	8.0
Platinum	9.3	Porcelain	3 to 4
Zinc	30.0	Fused quartz	0.56
Brass	18.4	Ice	52.0
Nickel-steel (36 % Ni) . .	1.0	Hard rubber	80.0

155. Practical Importance of Expansion. The forces exerted by bodies expanding or contracting under change of temperature are very great. Allowance must therefore be made for expansion in laying railroad rails or in designing girders for bridges. In a system of steam or hot water pipes expansion joints must be inserted. In riveting iron plates the rivets are placed in the holes while red-hot and hammered into shape

before they cool appreciably; on cooling, the rivets grip the plates together with a great force. Carriage tires are put on the wheel while hot, and the glass stopper of a flask when "stuck," may often be loosened by gently heating the neck of the flask.

Glass vessels are easily broken when suddenly heated or cooled, because the brittle glass cannot support the internal strains produced, before the temperatures of the outer and inner portions have become equalized. Porcelain will stand sudden changes much better, and fused quartz, whose coefficient of linear expansion is only $\frac{1}{14}$ of that of glass, may even be heated red-hot and plunged immediately into cold water without being broken. Invar steel is an alloy of nickel and iron with very small temperature coefficient and is used extensively for the construction of pendulums, steel tapes, etc.

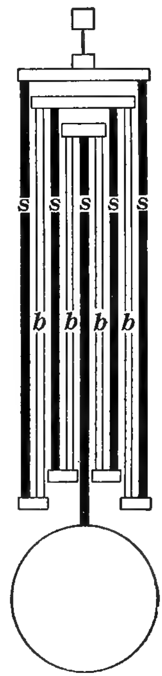


FIG. 83.

156. Further Applications. Any variation in the length of the pendulum will change the rate of a pendulum clock. Several devices have been employed to maintain the length constant with changing temperature. In the gridiron pendulum (Fig. 83) the lengthening of the rods marked s, usually of steel, lowers the bob, while the expansion of the rods b, usually of brass, raises the bob. If the total effective lengths of the two systems of rods be called L_s and L_b, and α_s and α_b be their coefficients of linear expansion, the length of the pendulum remains the same with change of temperature, if the expansion of the brass be made equal to the expansion of the steel, or if

$$\alpha_s L_s(t_2 - t_1) = \alpha_b L_b(t_2 - t_1) \qquad (221)$$

This condition is fulfilled when

$$L_s/L_b = \alpha_b/\alpha_s \qquad (222)$$

In the mercury compensating pendulum (Fig. 84) the lengthening or shortening of the rod is counteracted by a rise

or fall of the center of gravity due to the expansion of mercury, either in a vessel forming the bob of the pendulum or inclosed in the hollow stem of the pendulum rod.

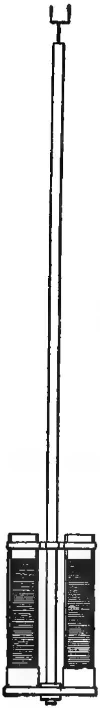


FIG. 84.

In chronometers and the better grade of watches the rate is kept constant by making the rim of the balance wheel of two different metals, the one with the larger coefficient being on the outside (Fig. 85). If the temperature rise, the rim will bend so as to decrease the diameter and consequently the moment of inertia, producing a more rapid motion of the wheel. This is made to balance exactly the effect of temperature upon the elasticity of the spring and the effect of the expansion of the diameter, both of which would result in a slower motion of the wheel.

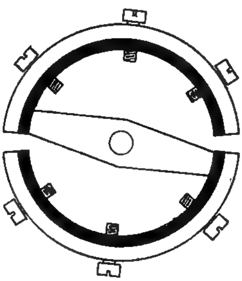


FIG. 85.

The same principle is employed in the so-called strain thermometers or "metallic" thermometers, frequently used in self-recording instruments. Their action may easily be understood from Fig. 86.

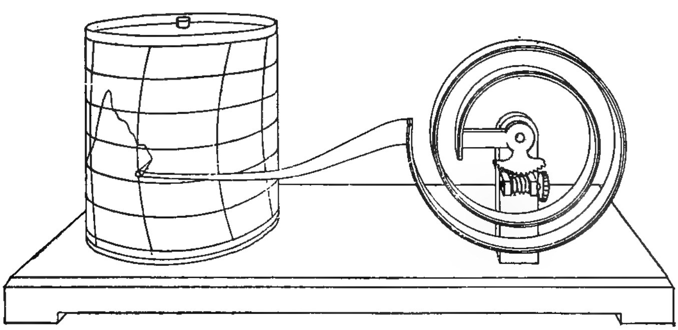


FIG. 86.

157. Cubical Expansion of Solids. An isotropic body is one whose physical properties are the same in all directions. Con-

sequently an isotropic body when heated, expands uniformly in all directions. Let us consider a parallelopiped cut from such a body. Let the three dimensions of this parallelopiped, at 0° C, be L'_0, L''_0, and L'''_0; at $t°$ C these dimensions will have become

$$\left.\begin{aligned} L'_t &= L'_0\,(1+\alpha t) \\ L''_t &= L''_0\,(1+\alpha t) \\ L'''_t &= L'''_0\,(1+\alpha t) \end{aligned}\right\} \tag{223}$$

and

$$L'_t \cdot L''_t \cdot L'''_t = L'_0 \cdot L''_0 \cdot L'''_0\,(1+\alpha t)^3 \tag{224}$$

$$= L'_0 \cdot L''_0 \cdot L'''_0\,(1+3\,\alpha t) \tag{225}$$

since we may neglect terms containing α^2 and α^3 in comparison with $3\,\alpha$.

Also since

$$V_0 = L'_0 \cdot L''_0 \cdot L'''_0 \tag{226}$$

and

$$V_t = L'_t \cdot L''_t \cdot L'''_t \tag{227}$$

we have

$$\left.\begin{aligned} V_t &= V_0\,(1+3\,\alpha t) \\ &= V_0\,(1+\beta t) \end{aligned}\right\} \tag{228}$$

where

$$\beta = 3\,\alpha \tag{229}$$

The quantity β is called the *coefficient of cubical expansion.*

In cubical as in linear expansion there exists no strict proportionality between increase of volume and increase of temperature. This is the reason why the readings of thermometers filled with different liquids, such as mercury and alcohol, do not exactly agree. If an empty flask be heated, it expands as if it were solid throughout, and the internal cavity increases by the same amount as would a solid body of the same form as the cavity and made of the same material as the flask.

* **158. Anomalous Expansion.** Most substances expand upon being heated. There are, however, some exceptions to the rule such as iodide of silver, between $-60°$ and $+142°$, cuprous oxide, below 4°, diamond, below $-39°$, and fused quartz, below $-80°$. The most important exception is water (Art. 160).

A rubber tube, when stretched to twice its original length or more, contracts when steam is passed through it. Accurate measurements, however, show that its volume increases. The

rubber has become *anisotropic* and its expansion in different directions is different, just as in all crystals, except those of the regular system.

159. Expansion of Liquids.[1] The cubical expansion of liquids and gases is much larger than that of solids. If a vessel containing a liquid be heated, both liquid and vessel expand, and the rise of the liquid in the vessel, as, for instance, the rise of mercury in a thermometer, indicates merely the *relative* expansion of the two substances. The effect of the expansion of the vessel is easily shown by filling a flask completely with a colored liquid, and closing it by a stopper through which a narrow tube passes, so that the liquid stands at a certain height in the tube. If now the flask be placed in warm water, the liquid will be seen at first to sink in the tube and then to rise, owing to the fact that the flask is heated first. Owing to the expansion of the vessel, the rise of the liquid in the tube is less than it would be if the vessel did not expand. The real coefficient of cubical expansion of the liquid, β, is therefore the coefficient of the vessel, β_v, plus the apparent coefficient β_a of the liquid, or

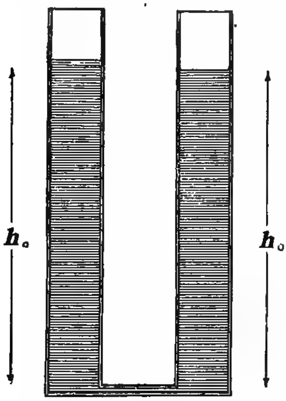


FIG. 87.

$$\beta = \beta_v + \beta_a \tag{230}$$

The real coefficient of expansion of the liquid may be determined directly by heating to different temperatures the two arms of a U-tube (Fig. 87) partly filled with the liquid in question. The pressures due to the columns on both sides balance in the horizontal tube. Let h_0 be the height of the meniscus above the horizontal tube at the lower temperature t_0, and let the density of the liquid on this side be d_0; let h_t, t, and d_t be the corresponding values on the other side. Then

$$h_0 d_0 g = h_t d_t g \tag{231}$$

[1] *For the determination of the coefficient of expansion of a liquid, see Manual, Exercise 39.*

But for a given mass M of the liquid

$$M = V_0 d_0 = V_t d_t \tag{232}$$

or

$$\frac{V_t}{V_0} = \frac{d_0}{d_t} = \frac{h_t}{h_0} \tag{233}$$

Now from (228) $\quad V_t = V_0[1 + \beta(t - t_0)]$ (234)

Therefore, if t_0 be chosen as 0° C,

$$\frac{h_t}{h_0} = \frac{d_0}{d_t} = \frac{V_t}{V_0} = 1 + \beta t \tag{235}$$

and

$$h_t = h_0(1 + \beta t) \tag{236}$$

from which β can easily be found. In this way the coefficient of expansion of mercury has been found to be 0.0001818 per degree.

The coefficient β increases considerably with increasing temperature and becomes quite large near the boiling point. Equation (234) should therefore be considered only as an approximation and, in the case of liquids, it is better to write

$$V_t = V_0(1 + \beta' t + \beta'' t^2) \tag{237}$$

Thus for alcohol $\beta' = 1020 \times 10^{-6}$ and $\beta'' = 200 \times 10^{-8}$,
for ether $\beta' = 1480 \times 10^{-6}$ and $\beta'' = 350 \times 10^{-8}$.

160. Maximum Density of Water. Water under 4° C forms an important exception to the general rule that bodies expand upon being heated. When cooled, under atmospheric pressure, from higher temperatures, it contracts, reaches the smallest volume at 4° C, and expands again upon further cooling. Figure 88 shows the variation of its specific volume (Art. 5), with temperature. If freezing be prevented by avoiding mechanical disturbances (Art. 195), the expansion continues below 0° C, as is indicated by the dotted line.

When a large body of water, as a lake, is cooled, the cooler, upper layer sinks to the bottom until a temperature of 4°C is reached. After this the cold water remains at the surface and it is there that the formation of ice begins, while the water at greater depths is still at 4°C. The temperature of maximum

density is considerably lowered by salts dissolved in the water. Thus sea water freezes before it reaches its maximum density. The temperature of the ocean at great depths is 2°.5 C.

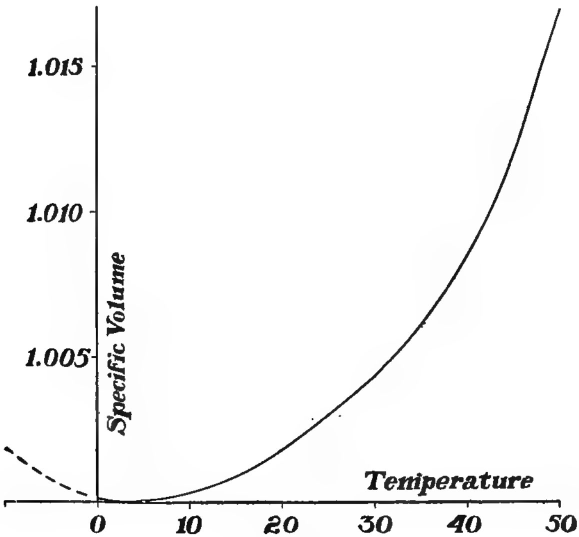


FIG. 88.

161. Expansion of Gases. Law of Gay-Lussac. In 1802 Gay-Lussac discovered that gas, when heated under constant pressure, expands equally for all equal differences of temperature, and that all gases possess the same coefficient of cubical expansion. This is known as the law of Gay-Lussac,[1] although it is frequently referred to as the law of Charles. The law is by no means exact. More recent investigations have shown that the coefficient of expansion is considerably larger in the case of gases which may be readily liquefied, such as chlorine and sulphur dioxide, than in the case of air, hydrogen or oxygen (Table VII). For the latter gases the coefficient of expansion has the nearly constant value 0.00367 per degree. This is considerably larger than the coefficient of expansion of solids or liquids under ordinary conditions. If

[1] Gay-Lussac, *Ann. Chim. et Phys.*, 1st ser., vol. 43.

we denote the coefficient of expansion of a gas under constant pressure by α_p, the general equation (235) may be written

$$V_t = V_0(1 + \alpha_p t) \tag{238}$$

and

$$d_t = \frac{d_0}{1 + \alpha_p t} \tag{239}$$

162. The Constant Pressure Gas Thermometer. Owing to its large coefficient of expansion, an inclosed mass of gas forms an extremely sensitive indicator of temperature changes, and so-called gas thermometers were used long before the appearance of mercury-in-glass thermometers. The air thermometer (Fig. 89) was probably invented by Galileo. In its simplest form it consists of a small bulb sealed to a long tube, which is mounted vertically with its lower end immersed in a suitable vessel filled with colored liquid. By expulsion of a few bubbles of air the liquid mounts into the tube. If the temperature of the bulb be varied slightly by placing the hand upon it, the colored index is promptly displaced through a considerable distance.

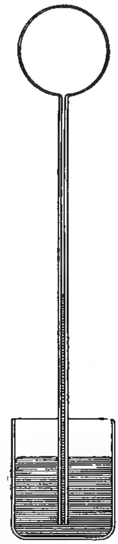


Fig. 89.

If we agree that changes in temperature, corresponding to equal changes in volume of a given mass of gas, under constant pressure, shall be called equal changes in temperature, we shall get a new temperature scale, but we should hardly expect this new scale to agree exactly with the one defined in Art. 149. The constant pressure gas thermometer is at present but little used, except as an indicator of temperature changes. The particular form known as the "Differential thermometer" (Fig. 90) may be made extremely sensitive, showing large displacements of the liquid index in the horizontal tube for slight differences of temperature between the two bulbs containing the gas.

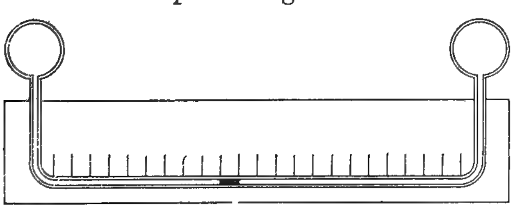


Fig. 90.

CHAPTER XX

INFLUENCE OF TEMPERATURE UPON THE PRESSURE OF A GAS

163. The Pressure Coefficient. When a mass of gas inclosed in a rigid vessel is heated, the increased kinetic energy of the gas produces an increased pressure in the vessel. When the volume is kept constant, the increase in the pressure exerted by the gas is very nearly proportional to the pressure and to the increase in temperature, as indicated by the mercury-in-glass thermometer,

or
$$P_t = P_0(1 + \alpha_v t) \tag{240}$$

where α_v is called the pressure coefficient of the gas, and the subscript v denotes that the volume is to remain constant. For perfect gases, namely, those for which Boyle's law holds, the relation between the pressure coefficient and the coefficient of expansion is readily found. For such gases we have by equation (238)

$$V_t = V_0(1 + \alpha_p t)$$

Consider such a gas at 0° C, having a volume V_0 under a pressure P_0. Let the temperature be changed to $t°$, *keeping the pressure constant.* Then the product of pressure and volume at temperature $t°$ is

$$P_0 V_t = P_0 V_0\ (1 + \alpha_p t) \tag{241}$$

Take the same volume of gas at 0° C, and raise the temperature again to $t°$, but this time *keeping the volume constant.* Then the product of pressure and volume at the same temperature, $t°$, is

$$P_t V_0 = P_0 V_0\ (1 + \alpha_v t) \tag{242}$$

In both cases the gas is finally at $t°$, and according to Boyle's law

$$P_0 V_t = P_t V_0 = PV \tag{243}$$

This in connection with equations (241) and (242) shows that

$$\alpha_p = \alpha_v = \alpha \tag{244}$$

Hence for perfect gases the two coefficients are equal and we may write:

$$PV = P_0 V_0 (1 + \alpha t) \tag{245}$$

which is known as the Boyle Gay-Lussac law.

The following table shows that hydrogen, oxygen, nitrogen and air may under ordinary conditions be considered as very nearly perfect gases. Others, which may be readily liquefied, may be treated as perfect gases only at high temperatures which are considerably above their points of liquefaction.

TABLE VII

COEFFICIENTS OF EXPANSION AND PRESSURE FOR GASES REFERRED TO VOLUME AND PRESSURE AT 0° C

SUBSTANCE	α_p $P = 76$ cm of mercury	α_v $P_0 = 76$ cm of mercury
Hydrogen	0.003660 per degree	0.003663 per degree
Oxygen	3674 " "	3665 " "
Nitrogen	3671 " "	3668 " "
Air	3671 " "	3665 " "
Carbon dioxide	3710 " "	3690 " "
Sulphur dioxide	3903 " "	3670 " "

164. The Constant Volume Gas Thermometer. If the volume of a given mass of gas be kept constant, equal to V_0, equation (245) becomes

$$P_t = P_0 (1 + \alpha t) \tag{246}$$

which, as previously shown, equation (240), is very nearly true for temperatures measured by the mercury-in-glass thermometer. This equation may be used as the basis of a new temperature scale. In order to do so we must make two fundamental assumptions.

(*a*) *Differences in temperature are to be taken as strictly pro*

portional to those differences in pressure which they produce in a mass of gas, maintained at constant volume.

(*b*) *The pressure P_0 is the pressure, exerted by this same mass of gas, at the temperature to be chosen as the zero of the new scale.* This will make the differences in temperature inversely proportional to the original pressure P_0, equation (249). If, now, we make our temperature readings to depend upon these assumptions, we see at once that equation (246) is made rigorously true. We should naturally expect any indicated temperature θ as read from this scale to be somewhat different from that given by the mercury-in-glass thermometer, which we denote by t. Then for a given temperature range t, read off on the mercury-in-glass thermometer, we have, substituting θ for t.

$$P = P_0(1 + \alpha\theta) \tag{247}$$

$$P - P_0 = \alpha P_0 \theta \tag{248}$$

and, finally,

$$\theta = \frac{1}{\alpha} \cdot \frac{P - P_0}{P_0} \tag{249}$$

If, now, we call as before the freezing point and boiling point of water 0° and 100°, and denote the pressure at the latter temperature by P_{100}, we have

$$P_{100} - P_0 = \alpha P_0 100 \tag{250}$$

and

$$\theta = 100 \frac{P - P_0}{P_{100} - P_0} \tag{251}$$

To find the temperature θ, therefore, it is only necessary to measure the pressures at the fundamental points and at the temperature to be determined.

In its simplest form (Fig. 91) this type of gas thermometer consists of a bulb, A, containing the gas, connected by a capillary tube to a U-tube filled with mercury. In order to insure constancy of volume of the inclosed gas, the mercury is adjusted for each reading, so that it just touches a pointer or mark, p, placed near the opening of the capillary tube, or on the tube itself. The pressures are calculated from the differ-

ence in height of the mercury in the two arms of the U-tube in each case. To the temperature θ, as computed from equation (251), there must be applied certain corrections, due to the expansion of the bulb, changes in barometric pressure, etc.[1]

165. The Standard Hydrogen Thermometer. Owing to certain serious limitations (Art. 150), the mercury-in-glass thermometer is not well adapted to serve as a final standard for measurements of temperature. In its stead the constant volume gas thermometer has, by international agreement, been adopted as the standard instrument for temperature measurements. Further, since hydrogen fulfills most nearly the requirements for a perfect gas, it has been adopted as the thermometric substance, with the provision that the gas at freezing point shall be under a pressure of 100 cm of mercury.

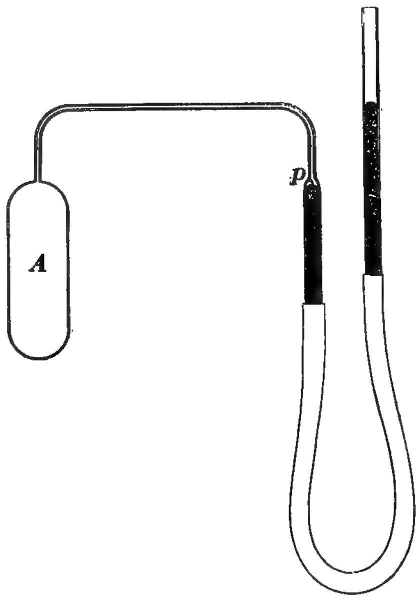


FIG. 91.

Temperatures measured by this thermometer are the final standards in all scientific work, and temperatures measured by any other means are usually "reduced to the hydrogen scale" when any high degree of precision is required. It should be said, however, that in the case of mercury-in-glass thermometers from the best makers, the readings do not differ greatly from those of the hydrogen scale, while the greater convenience of the mercury instruments more than compensates for its limitations, except in cases where extreme accuracy is required.

[1] *See Manual, Exercise 40.*

Throughout the remainder of this text, distinctions between temperatures based upon the mercury or the hydrogen scale will be disregarded. It is also to be understood that, in general, temperatures will be expressed in Centigrade degrees, unless it be expressly stated to the contrary.

166. The Zero of the Gas Scale. Absolute Temperatures. The selection of the freezing point of water as the zero point was an entirely arbitrary one. It is now possible to find a more rational zero point. From the principles developed in the discussion of the constant volume thermometer (Art. 164) it is clear that a decrease in temperature is indicated by a corresponding decrease in gas pressure. The rational zero for such a temperature scale would therefore be that temperature for which the pressure of the gas in the hydrogen thermometer becomes zero, assuming that hydrogen could remain a gas at that temperature. This zero point may readily be found from the equation

$$P_t = 0 = P_0 (1 + \alpha t) \tag{252}$$

Whence $$t = -\frac{1}{\alpha} = -\frac{1}{0.00366} = -273° \text{ nearly.}$$

Thus we see that this zero is about 273° below the zero of the Centigrade scale. This point is called the zero point of the hydrogen scale, or the *absolute zero*. Temperatures measured on this scale are called *absolute temperatures*. Such readings are usually denoted by T, and are related to readings t, on the Centigrade scale through the equation

$$T = t + 273° \tag{253}$$

The terms *absolute zero* and *absolute temperatures* are not well chosen, inasmuch as this zero and all temperatures measured from it depend upon the use of some arbitrary thermometric substance.

In advanced physics the student will become acquainted with a temperature scale, called the thermodynamic scale, in which the temperatures are independent of any particular thermometric substance. Temperatures, expressed in this scale, are

"absolute" temperatures. They agree very closely with the absolute temperatures derived from the gas scale, and it seems unnecessary to distinguish between the two in this text. For a discussion of the thermodynamic scale the student is referred to more advanced textbooks.

In accordance with the molecular theory of heat, the pressure of an inclosed mass of gas is due to the bombardment of the sides of the vessel by the flying molecules. From this it follows that the pressure decreases as the molecular motion decreases, and consequently (Art. 147) at the absolute zero all molecular motion would cease. Hydrogen is not a perfect gas. It is due to this fact that the two sets of "absolute" temperatures referred to above do not exactly coincide.

167. The Gas Law. For a perfect gas at temperature $t°$ we have found

$$PV = P_0 V_0 (1 + \alpha t) \tag{254}$$

Introducing the absolute temperature, we have

$$1 + \alpha t = \alpha\left(\frac{1}{\alpha} + t\right) = \alpha(273 + t) = \alpha T \tag{255}$$

$$PV = \alpha P_0 V_0 T = RT \tag{256}$$

where R is a constant for a given mass of gas and may be called the *gas constant.*

For any other absolute temperature T',

$$P'V' = RT' \tag{257}$$

By division of the last two equations we obtain the important relation

$$\frac{PV}{P'V'} = \frac{T}{T'} \tag{258}$$

This is the mathematical statement of what is called the *gas law.* In words, it says that "*For a perfect gas the product of pressure and volume is proportional to the absolute temperature.*"

From equation (258) follow directly the two corollaries: "The pressure of a gas of constant volume varies directly as

the absolute temperature" and "The volume of a gas under constant pressure varies directly as the absolute temperature."

Since ordinary gases at temperatures and pressures remote from their points of liquefaction behave nearly as perfect gases, these laws may be applied to such gases without introducing any serious error.

Problems

1. Reduce the following readings on a Fahrenheit thermometer to the corresponding readings on a Centigrade thermometer: $+2000°$, $+104°$, $+5°$, $-49°$. *Ans.* $+1093°.3$; $+40°$; $-15°$; $-45°$.

2. Reduce the following readings on a Centigrade thermometer to the corresponding readings on the Fahrenheit thermometer: $+2000°$, $+445°$, $-10°$, $-273°$. *Ans.* $+3632°$; $+833°$; $+14°$; $-459°.4$.

3. At what temperature do Fahrenheit and Centigrade thermometers read the same? *Ans.* $-40°$.

4. What are the coefficients of linear expansion of iron and copper when expressed in English units?

Ans. Iron 6.722×10^{-6}; copper 9.5×10^{-6} per degree F.

5. A certain railroad track is laid when the temperature is 32° F. If the rails are 30 ft long, how much space in inches must be left between them in order that they may just touch when the temperature is 122° F?

Ans. 0.218 in.

6. The steel cable from which the Brooklyn bridge hangs is more than a mile long. How many feet does a mile of its length vary between a winter day when the temperature is $-20°$ C and a summer day when it is 30° C? ($a = 12.1 \times 10^{-6}$.) *Ans.* 3.19 ft.

7. In a gridiron pendulum (Fig. 83), the distance from the center of suspension to the center of oscillation is 99.3 cm. Supposing the whole pendulum except the vertical brass rods to be made of iron, how long must the brass rods be, for perfect compensation, all brass rods being the same length?

Ans. 95.36 cm.

8. A clock with a seconds pendulum of iron is correct at 25° C. How many seconds per day will it gain if the temperature is 0°? What will be the result if the pendulum be made of nickel-steel?

Ans. 13.07 sec; 1.08 sec.

9. How large a relative error is made if the moment of inertia of a brass cylinder, calculated from measurements at 10°, should be used in experiments carried on at 50°? *Ans.* 0.147 per cent.

10. A certain glass measuring flask made of Jena glass has a capacity of one liter at 15°. What is its capacity at 40°? *Ans.* 1000.6 cm^3.

11. The density of mercury at 0° is 13.596 g/cm^3. What is its specific volume at 35°? *Ans.* 0.07402 cm^3 per gram.

12. Reduce to 0° a barometer reading of 74.5 cm taken at 19°, correcting both for the expansion of the mercury and for that of the brass scale. *Ans.* 74.27 cm.

13. A piece of brass weighs 17 g in air and 15 g in water at 10°. How much will it weigh in water at 30°? Spec. vol. of water at 0° C = 1.00012, at 10° C = 1.00027, at 30° C = 1.00435 cm^3 per gram. *Ans.* 15.00592 g.

14. A liter flask of Jena thermometer glass is calibrated at 0°. The cylindrical neck has a diameter of 1.5 cm. How far above the liter mark will the meniscus be, if the flask be filled with one kilogram of water at 30°? *Ans.* 2.06 cm.

15. A quantity of air, at atmospheric pressure and 0°, occupies 1000 cm^3. It is heated to 100°, under constant pressure. Find the increase in volume and the work done during the expansion. *Ans.* (*a*) 367.1 cm^3.
(*b*) 37.17 joules.

16. What fractional part of the air in a room passes out when the air is heated from $-10°$ to 20°? *Ans.* 0.11.

17. The density of air under standard conditions is 0.00129 g/cm^3. What will be the mass of air in a room whose dimensions are $8 \times 5 \times 3$ m, the temperature being 20° and the barometric reading 730 mm? *Ans.* 138.86 kilos.

18. If the atmospheric pressure on a mountain top (Mont Blanc) be 380 mm of mercury, and the temperature $-10°$ C, what fraction of the normal quantity of air is contained in any given volume? *Ans.* 0.519.

19. Suppose 20 cc of hydrogen gas to have been collected in a closed tube over mercury. The barometer reading corrected for temperature is 74.5 cm and the temperature of the gas is 24°. The mercury in the tube stands 12 cm above the mercury in the dish in which the tube is inverted. Reduce the volume of the gas to standard conditions. *Ans.* 15.12 cm^3.

20. A certain boiler boils dry, containing steam at 80 pounds per square inch, at 155° C, which in the absence of water may be considered as a perfect gas. If the temperature should rise to 600° C, would the boiler burst, if it can just support a pressure of 300 pounds per square inch? *Ans.* No. $P = 163.2$ lb per in^2.

QUANTITY OF HEAT

CHAPTER XXI

CALORIMETRY

168. The Unit of Heat. The Calorie. In calorimetry we are concerned with the measurement of definite quantities of heat. The idea of quantity of heat involves first the mass of the body to be heated. We know that the temperature of boiling water is the same, whether the vessel contain one liter or ten liters, and yet it will take a flame ten times as long to heat ten liters of water from zero to the boiling point, as to do the same for one liter, or (Art. 146) the heat required is ten times as great in the first case as in the second. Hence the quantity of heat needed to raise the temperature of a body through a given temperature interval is proportional to the mass of the body.

Again, we assume that the heat absorbed by a body when heated to a certain temperature is all given out when the body cools to its original temperature. Consequently when two kilos of water at 20° are mixed with two kilos of water at 0°, we expect the mixture to assume a final temperature of 10°, and experiment shows this to be true (Art. 173). If, on the other hand, we should substitute two kilos of iron or two kilos of mercury for the two kilos of water at 20° in the experiment, we should find that equal masses of different substances give out very different quantities of heat when cooled through the same temperature range. Thus a mass of mercury absorbs or gives out only *one thirtieth* as much heat as an equal mass of water would do when carried through the same temperature interval.

It is clear, then, that the idea of the quantity of heat needed to heat a body involves three distinct things: (*a*) the mass of the body, (*b*) the material of which the body is composed, and (*c*) the temperature range through which it is to be heated. It is therefore necessary to specify all these things in defining the unit of heat. The accepted unit of heat is defined as *that quantity of heat which will raise the temperature of one gram of water from 15 to 16 degrees Centigrade. This unit is called a calorie.*

In the English system the British thermal unit (B. T. U.) is defined as the quantity of heat needed to raise the temperature of one pound of water one degree Fahrenheit.

169. Thermal Capacity of a Body. It may be proven experimentally (Art. 173) that, unless changes of state occur, the quantity of heat added to a body is nearly proportional to the change of temperature produced. Thus if a given *body* be heated from $t_1°$ to $t_2°$, then the heat needed for this purpose is

$$H = C(t_2 - t_1) \tag{259}$$

The proportionality factor,

$$C = \frac{H}{t_2 - t_1} \tag{260}$$

is called the *thermal capacity of the body* and its unit is one calorie per degree.

170. Thermal Capacity of a Substance. It has likewise been found that the heat needed to produce a definite temperature change $(t_2 - t_1)$ in different masses of the same substance is proportional to the mass, or

$$H = cM(t_2 - t_1) \tag{261}$$

The proportionality factor

$$c = \frac{C}{M} \tag{262}$$

which is evidently characteristic of the substance, is called the *thermal capacity of the substance*, and is *numerically* equal

to the heat necessary to change the temperature of *one gram of the substance* one degree. Its unit is one calorie per gram-degree.

171. Thermal Capacity of Water. The above formulae are not quite exact, since in general the thermal capacity of a substance increases with the temperature. Since water has been chosen as the standard substance, we have numerically

$$\left.\begin{aligned} c_w &= 1 \\ M &= 1 \\ t_2 - t_1 &= 1 \end{aligned}\right\} \tag{263}$$

Consequently the change of the thermal capacity of water is of especial interest. In Fig. 92 we have the thermal capacity of water plotted as a function of the temperature.

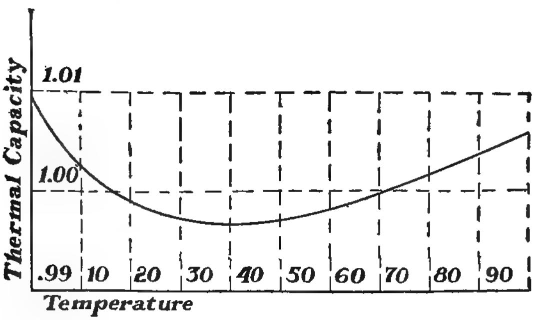


FIG. 92.

Though these variations are small, yet it is evident that a definite temperature must be specified in the definition of the calorie. The interval from 15° to 16° has been chosen since the calorie based upon this interval is the one one-hundredth part of the heat needed to raise the temperature of one gram of water from 0° to 100°. In this text it will be assumed that the thermal capacity of water c_w may always be taken as unity.

172. Specific Heat of a Substance. The specific heat s of a

substance is *the ratio of the thermal capacity of the substance to that of water*, or

$$s = \frac{c}{c_w} \tag{264}$$

Since c_w is unity, the specific heat of a substance is numerically equal to its thermal capacity. In a loose sense the thermal capacity is often called specific heat, but the student should observe that the same relation holds between these two quantities as between density and specific gravity. The former is a definite physical quantity; the latter is simply a ratio or a pure number.

173. The Method of Mixtures.[1] All heat measurements are based upon the following fundamental principle: *When two or more bodies, originally at different temperatures, are placed in thermal contact, and exchange of heat takes place exclusively between these bodies, the heat lost by one part of the system is equal to the heat gained by the other.* This is called the principle of equal heat exchanges.

Thus if a certain mass M of a substance, whose thermal capacity is c_1, be heated to a temperature $t_1°$ and then dropped into a mass of water M_w, at temperature t_2, where $t_2 < t_1$, an exchange of heat takes place and an intermediate temperature t is established. The body loses an amount of heat equal to $c_1M_1(t_1 - t)$, and the water gains an amount equal to $c_wM_w(t - t_2)$. Consequently we have

$$c_1M_1(t_1 - t) = c_wM_w(t - t_2) \tag{265}$$

and the specific heat of the substance is

$$s = \frac{c_1}{c_w} = \frac{M_w(t - t_2)}{M_1(t_1 - t)} \tag{266}$$

The water is contained in a vessel, supplied with a stirrer and a thermometer. These form the *calorimeter*. Their temperature is changed as well as that of the water. The heat given to them is equal to the sum of such terms as $cM(t - t_2)$, calculated for all parts of the calorimeter. The sum of the products

[1] *For determination of specific heat by method of mixtures, see Manual, Exercise 42.*

of the specific heats into the mass $\Sigma \frac{c}{c_w} M$ is called the *water equivalent* of the calorimeter, and must be added to the mass of water M_w, in the experiment just described. Taking into account the heat effect of the calorimeter, we obtain then the general equation:

$$s = \frac{c_1}{c_w} = \frac{\left(M_w + \Sigma \frac{c}{c_w} M\right)(t - t_2)}{M_1(t_1 - t)} \tag{267}$$

TABLE VIII

SPECIFIC HEATS OF SOLIDS AND LIQUIDS

SUBSTANCE	s	ATOMIC W.	s × AT. W.
Copper	0.094	63.6	5.98
Graphite	0.199	12.0	2.39
Iron	0.116	55.9	6.48
Lead	0.031	206.9	6.42
Mercury	0.033	200.0	6.60
Platinum	0.032	194.8	6.23
Tin	0.055	119.0	6.55
Zinc	0.094	65.4	6.15
Glass	0.200	—	—
Ice	0.505	—	—
Alcohol	0.602	—	—
Ether	0.547	—	—

***174. Law of Dulong and Petit.** The above table shows that the specific heats of different substances vary considerably. In 1819 Dulong and Petit announced the following law:[1] "The product of the specific heat of a substance into its atomic weight is the same for all elementary solid substances." The product thus obtained is about 6.

This law is by no means exact. Carbon, boron and silicon are exceptions. Since the specific heats vary with the temperature, these products will vary according to the temperature chosen.

[1] Dulong et Petit, *Ann. Chim. et Phys.*, 1819.

Nevertheless this law suggests the possibility that the thermal capacities of atoms of different substances may be nearly the same.

For a more complete discussion of this law and others relating to the molecular heats of chemical compounds the student is referred to textbooks on physical chemistry.

175. Specific Heats of Gases. Very different values may be obtained for the specific heat of a gas, according to the conditions under which the gas is heated. If a gas be *allowed to expand* while it is being heated, it will do work against external pressure by virtue of that expansion. The energy needed to do this work must be supplied from the gas itself in the form of heat. Consequently the gas heated under this condition absorbs heat for two reasons: (*a*) *to produce change* of temperature, (*b*) *to furnish energy* to do the work of expansion.

On the other hand if the gas be heated and its *volume be kept* constant, no heat is absorbed except that needed to produce the rise in temperature. The two specific heats obtained under these two conditions are designated s_p, *the specific heat under constant pressure*, and s_v, *the specific heat under constant volume.* Of these two values, s_p is greater, by the amount of heat needed to produce the expansion. The ratio between these two values, for atmospheric pressure, usually denoted by γ, is of considerable importance in the computation of the velocity of sound in a gas (Art. 113).

TABLE IX

SPECIFIC HEATS OF GASES

SUBSTANCE	s_p	s_v	γ
Air	0.237	0.167	1.41
Hydrogen	3.410	2.418	1.41
Oxygen	0.217	0.147	1.41
Water vapor	0.421	—	1.30
Carbon dioxide	0.203	—	1.30
Alcohol	0.453	—	1.13

CHAPTER XXII

THE MECHANICAL THEORY OF HEAT

176. The Experiments of Joule and Rowland. We have seen (Art. 146) that heat must be considered as a form of energy and that heat is produced whenever mechanical energy is absorbed in overcoming friction. The first accurate experiments showing a definite quantitative relation between heat and mechanical work were made by Joule[1] (1818–1889). In his apparatus (1845) descending weights M were made to rotate a paddle wheel P immersed in a stationary calorimeter C (Fig. 93). The resistance offered to the motion of the paddle was greatly increased by means of stationary vanes extending into the interior of the calorimeter. The work done by the total mass $2M$, descending through a height h, is $2Mgh$. The heat produced, measured in heat units, is

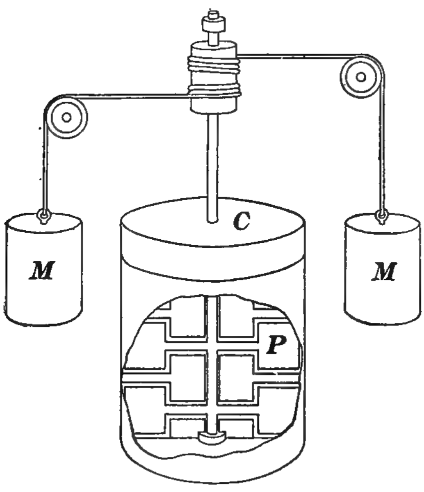


FIG. 93.

$$H = (t_2 - t_1)\Sigma cm \tag{268}$$

where c and m denote the thermal capacities and the masses of the different parts of the calorimeter and $t_2 - t_1$ the corresponding rise of temperature. Joule showed that for every unit of

[1] Joule, *Phil. Trans. Roy. Soc.*, 1850.

mechanical energy which disappears, the same quantity of heat is always produced.

In later experiments[1] (1878) he used a calorimeter which was free to turn about a vertical axis. If the paddle be turned, the friction in the water tends to rotate the calorimeter. The vessel, however, is kept stationary by means of two thin silk strings wound in a groove around the cylindrical vessel and leaving it in a tangential direction at two opposite ends of a diameter. These strings then pass over two light pulleys and carry at their lower ends weights M, which are adjusted until the calorimeter remains stationary, when the paddles revolve at a constant rate. If r be the radius of the groove, the moment of the mechanical forces counteracting the effect of the paddle wheel is $2\,Mgr$. The total work done in t seconds when the paddles are rotated n times per second against the torque due to the weights, is

$$W = torque \times angle = 2\,Mgr \times 2\,\pi nt = 4\,\pi ntrMg \qquad (269)$$

In 1879 Rowland (1848–1901) made a series of classical experiments[2] with an improved form of Joule's later apparatus, obtained much more accurate results and also discovered the variation of the thermal capacity of water with change of temperature (Art. 171).

177. The Mechanical Equivalent of Heat. Joule's and Rowland's experiments, as well as many others of more recent date, have shown that *in any transformation of heat into work or of work into heat, the ratio between the numerical values of these two forms of energy is always a constant.* If the number of heat units be denoted by H, and the number of units of work by W, then

$$W \text{ mechanical units} = JH \text{ heat units} \qquad (270)$$

If we combine equations (268), (269) and (270), we have

$$W = 4\,\pi ntrMg = J(t_2 - t_1)\,\Sigma cm \qquad (271)$$

or

$$J = \frac{4\,\pi ntrMg}{(t_2 - t_1)\,\Sigma cm}\,\frac{ergs}{cal} \qquad (272)$$

[1] Joule, *Phil. Trans. Roy. Soc.*, 1878.
[2] Rowland, *Proc. Am. Ac. Arts and Sci.*, 1879.

The value of J depends upon the units chosen.

Thus 778 foot-pounds of work are equivalent to one B. T. U., or 0.427 kilogram-meter to one calorie.

From a critical study of these experiments Barnes concluded that

$$\text{One calorie} = 4.186 \times 10^7 \text{ ergs} = 4.186 \text{ joules}$$

The mechanical work corresponding to one heat unit is called the *mechanical equivalent of heat.*

We have thus obtained a new measure of heat in terms of mechanical units, and a new measure of mechanical work in terms of heat units. Thus

$$H \text{ calories} = JH \text{ joules} = 4.186\ H \text{ joules}$$

and

$$W \text{ joules} = W/J \text{ calories} = 0.239\ W \text{ calorie}$$

178. The First Law of Thermodynamics. The experimental results given in the preceding paragraphs may be summarized thus: *When work is transformed into heat, or heat into work, the amount of work is always equivalent to the quantity of heat.* This is known as the first law of thermodynamics.[1]

Heat added to a body is considered as positive, heat given out by a body as negative; work done upon a body is positive; work done by the body, negative. Taking into account the signs, the mathematical expression for this law is:

$$W + H = 0 \tag{273}$$

when W and H are measured in the same units, or

$$W + JH = 0 \tag{274}$$

when W is expressed in mechanical, and H in heat units.

179. Equivalence of Energy and the Principle of Conservation. Having established the equivalence of mechanical work and heat, it is of the highest importance that we should grasp the full significance of this equivalence of energy which shows itself in every branch of Physics. Physical phenomena of every form

[1] *This law was first clearly stated by Mayer in Liebig's Annalen 1842.*

depend upon the transference or the transformation of energy. We are now prepared to say that energy in any of its manifold forms may be reduced to an equivalent amount of heat and hence to an equivalent amount of energy of any other form. Thus an electric current of strength I, flowing for t seconds through a resistance R, produces a quantity of heat H, which is always proportional to I^2Rt. This quantity, measured in electrical units, is called electrical energy. In short if, in any physical phenomenon, energy of any form disappear, energy of some other form will always appear, and the energy of the new form is always equivalent to the energy of the old. No energy is lost.

This most important principle, first announced by Robert Mayer in 1845, is known as the principle of conservation of energy. It is this principle, verified by countless experiments, which underlies all physical phenomena, and which constitutes one of the grandest generalizations of modern science. Keeping in mind that energy added to a body is positive, and energy taken away is negative, the principle may be stated in these words: If in a system of bodies no reaction be allowed between the system and outside bodies, then the total amount of energy of the system is not changed by any reaction or transformation between the parts of the system.

The actual amount of energy in a body is not known, but for any change from a state A to a state B, the energy involved can be accurately determined and is the same regardless of the manner in which the change has been accomplished. For, let us assume that a system of two bodies be changed from one state to another, during which change the first body gives to the second a certain amount of energy, and that the system can be brought back to its original state by a method in which the second body returns to the first a smaller amount of energy than it received from it. The result of the complete cycle of changes would be an increase of energy possessed by the second body without an equivalent compensation on the part of the first, and this contradicts the first law of thermodynamics.

By a bold generalization, which admits of no proof by actual

experiment, the principle is sometimes stated by saying that the energy of the universe is constant, or that energy can neither be created nor destroyed.

180. Compression and Rarefaction of a Gas. Let a cylinder, closed by a piston, contain a given amount of gas. If the piston be suddenly pushed in through a *small distance* s, with a force F, the work done upon the gas is (Art. 35)

$$W = Fs = Pv \tag{275}$$

where P is the pressure and v is the change of volume. This mechanical energy expended upon the gas increases its energy and consequently the gas is heated.

On the other hand, if the gas be allowed to do work against a pressure P, increasing its volume by v, a quantity of heat equivalent to the work Pv is abstracted from the gas and it cools.

181. Free Expansion of a Gas. Joule connected two receivers, A and B (Fig. 94), by a tube, containing a stopcock s. One of the vessels was exhausted while the other contained a gas under pressure. Both receivers were then immersed in a water calorimeter. Upon opening the stopcock the gas rushed into the vacuum and now filled both cylinders, but the calorimeter showed no change of temperature. From this Joule concluded that the energy of a given mass of gas is independent of the volume occupied.[1]

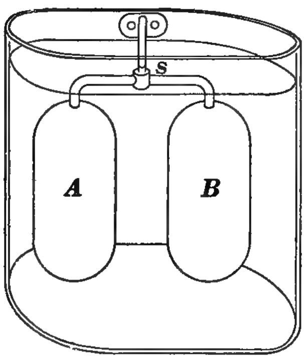


Fig. 94.

Placing each receiver in a separate calorimeter and repeating the experiment, the water in the vessel originally containing the compressed gas was cooled, and that in the other one was heated. Obviously the loss of energy on the one side equals the gain on the other.

[1] *This experiment was originally due to Gay-Lussac (Mem. d'Arcueil. 1807).*

The results obtained by Joule in this experiment hold rigorously only for perfect gases. Ordinary gases do show a small change of energy upon expansion, but this change must be measured by more sensitive methods (Art. 186).

182. Isothermal and Adiabatic Expansion. If a gas expand very slowly while kept in close thermal contact with a heat reservoir of temperature $t°$, heat will constantly flow from the reservoir to the gas and keep it at the temperature $t°$ during the expansion. Such a process is called an *isothermal expansion* and is represented by line I (Fig. 95) in which the volumes of the gas are plotted as abscissae, and the corresponding pressures as ordinates. The equation of this line (Art. 167) is

$$PV = RT = \text{constant}$$

The work done by the gas during a very small increase in volume v, equal to ab, is Pv, and is represented at the point Q of the curve in the figure, by the shaded area $Qcba$. Imagine a large number of such narrow strips drawn, side by side, and extending from the axis of volumes to the isothermal line. It is then clear that the work done by the gas during the isothermal expansion from P_0V_0 at point A to P_1V_1 at point B is equal to the sum of all the strips, and is therefore represented by the area $ABNL$ included between the isothermal line, the axis of volumes and the two ordinates representing P_1 and P_0.

If the gas, on the other hand, be inclosed in a vessel whose walls prevent any flow of heat through them, the gas can receive no heat from the reservoir and will therefore cool during expansion. The pressure P_2 or ND, corresponding to volume V_1 will be smaller than P_1 or NB, and the line representing this change in the gas will at the intersection A be *steeper than the isothermal line.* Such an expansion is called an *adiabatic expansion*, and is shown by the adiabatic line II (Fig. 95).

By the use of calculus it may be shown that the equation of an adiabatic line for a perfect gas is

$$PV^{\gamma} = \text{constant} \qquad (276)$$

where γ is the ratio of the two specific heats of the gas (Art. 175).

***183. Evaluation of $C_p - C_v$.** If a quantity of gas of mass M be heated under constant volume through a *temperature interval* of $t°$, the heat needed is Mc_vt calories. If it be heated under constant pressure, the heat needed is Mc_pt calories. According to Art. 181 the energy of a gas is independent of its volume, and consequently the quantity of heat needed for the mere heating of the gas must be the same in both cases. But, when heated under constant pressure, the gas does work equal to Pv, and this energy must also be supplied by the heat added to the gas, hence

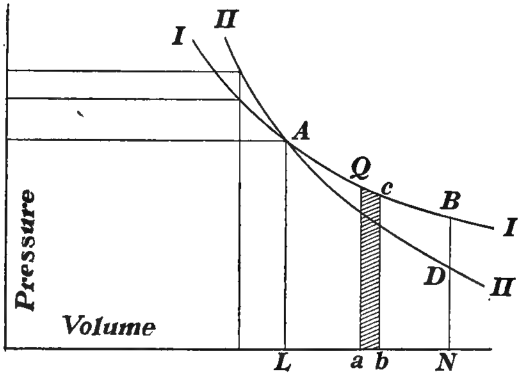


FIG. 95.

$$Mc_pt - Mc_vt = (c_p - c_v)\ Mt = Pv \tag{277}$$

According to Art. 163, equation 245, we have before heating

$$V = \frac{P_0 V_0}{P}(1 + \alpha t_1)$$

after heating

$$V + v = \frac{P_0 V_0}{P}(1 + \alpha t_2) \tag{278}$$

The increase in volume v is therefore

$$v = \frac{P_0 V_0}{P}\alpha(t_2 - t_1) = \frac{P_0 V_0}{P}\alpha t \tag{279}$$

and by (277) and (256)

$$c_p - c_v = \frac{P_0 V_0}{M}\alpha = \frac{R}{M}\ \frac{ergs}{gram\text{-}degree} \tag{280}$$

PV and R are usually given in mechanical units, hence, if we wish to express $c_p - c_v$ in thermal units, the right-hand member of the equation must be divided by J,

or $$c_p - c_v = \frac{P_0V_0}{JM}\alpha = \frac{P\alpha}{Jd_0} = \frac{R}{JM}\,\frac{calories}{gram\text{-}degree} \tag{281}$$

where d_0 is the density at 0° C.

Since the specific heats s_p and s_v are numerically equal to c_p and c_v,

$$s_p - s_v = \frac{P_0V_0}{JM}\alpha \text{ numerically} \tag{282}$$

***184. Coefficients of Volume Elasticity.** From Fig. 95 it appears that during an *adiabatic compression* the change in *pressure* must be *larger* to produce a given change of volume than in the case of an *isothermal compression.* From the definition of the coefficient of volume elasticity (Art. 58)

$$e = \frac{dp}{dV}\cdot V$$

it is seen that a distinction must be made between the adiabatic and isothermal coefficients, e_a and e_i. Let the changes in pressure corresponding to a small change of volume

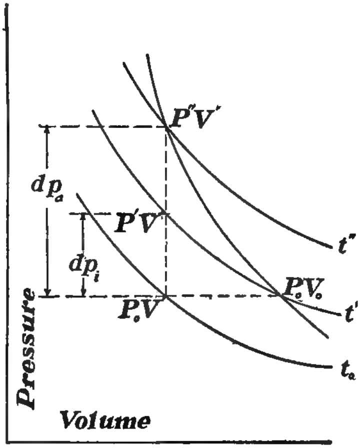


FIG. 96.

$$V_0 - V' = dV$$

be $$P'' - P_0 = dp_a$$

and $$P' - P_0 = dp_i$$

respectively (Fig. 96), then, as e_a and e_i vary as dp_a and dp respectively, dV being constant,

$$\frac{dp_a}{dp_i} = \frac{e_a}{e_i} \tag{283}$$

Let the corresponding changes of temperature $t'' - t_0$ and $t' - t_0$ be dt_a and dt_i. But dp_a and dp_i may also be considered as the increase in pressure of the gas when heated under constant volume V', through dt_a and dt_i degrees.

Since by Art. 167 (eq. 254),

$$P_0 + dp = \frac{P_0 V_0}{V} \cdot (1 + \alpha dt)$$

the increases in pressure, dp_a and dp_i, are proportional to the increases in temperature, dt_a and dt_i, or by (283)

$$\frac{e_a}{e_i} = \frac{dt_a}{dt_i} \tag{284}$$

Now consider the gas to be brought from $P_0 V'$ to $P'' V'$ by two different methods: (1) by heating under constant volume, for which the heat needed is $Mc_v dt_a$; (2) by heating under constant pressure, until its volume has become V_0, in which case the heat needed is $Mc_p dt_i$. Then compress the gas adiabatically until it reaches the state $V'P''$. No heat from the outside is needed for this. During the two parts of the last process, work is done by the gas during the expansion, and done *upon it* during compression; but if the increment of pressure be *very small* in comparison with the total pressure, the total external work may be considered as practically zero, in comparison with the heat involved.

The total changes in energy of the gas along the two paths may, therefore, be set equal to each other and equal to the heat absorbed, or

$$Mc_v dt_a = Mc_p dt_i \tag{285}$$

If we call the ratio of the two specific heats γ,

$$\frac{e_a}{e_i} = \frac{dt_a}{dt_i} = \frac{c_p}{c_v} = \frac{s_p}{s_v} = \gamma \tag{286}$$

***185. Velocity of Sound in a Gas.** As has been shown (Art. 111), the velocity V of sound in a gas is given by

$$V = \sqrt{\frac{e}{d}}$$

But since the compressions and rarefactions in the air occur under adiabatic conditions, the corresponding coefficient of volume elasticity must be used, or

$$e = e_a = \gamma e_i \tag{287}$$

and the expression for the velocity of sound becomes, since e_i is equal to the pressure P (Art. 113),

$$V = \sqrt{\gamma \frac{P}{d}}$$

The values for γ for several different gases under a pressure of one atmosphere are given in Table IX, on page 215.

***186. The Joule-Thomson Effect.** In the experiments described (Art. 181), Joule found no change of energy in a gas due to change of volume alone. If this were true, no forces would exist between the separate molecules of the gas, and the energy of the gas would simply be the sum of the kinetic energies of its molecules.

Joule and Thomson (Lord Kelvin) found, however, by their famous "porous plug" experiment,[1] that there is a slight change in energy when a gas expands. They passed gas slowly from a vessel at high pressure into the open air through a plug made of cotton wool. The temperature of the gas was measured just before entering and after leaving the plug, and it was found that in most gases the temperature after expansion was lower than before. In the case of hydrogen a heating was observed. It has since been shown, however, that even hydrogen gas will cool if the original temperature of the gas under high pressure be below $-80°$ C. The compressor producing the high pressure does work upon the gas equal to the product of the pressure of its piston into the change of volume. On the other hand, the gas at the lower pressure side does work equal to the product of its volume into this lower pressure. For a perfect gas these two amounts of work should be equal.

[1] Joule and Thomson, *Phil. Mag.*, 1852.

Of course we must consider the fact that Boyle's law does not hold strictly for ordinary gases, and that consequently the work Pv done upon the gas at the high pressure does not quite equal the work done by the gas on the low pressure side. However, after taking this into account, the porous plug effect, which in all cases is quite small, indicates that there is some intermolecular action in all ordinary gases, and that, in general, energy is required to produce a mere expansion.

CHAPTER XXIII

TRANSFORMATION OF HEAT INTO MECHANICAL ENERGY

187. Modes of Transformation. The transformation of mechanical energy or of any other form of energy into heat is of common occurrence and no difficulty is encountered in making such transformation complete. On the other hand, no method has ever been devised for reversing completely any process which has produced heat, although — according to the first law of thermodynamics — this would seem to be theoretically possible.

In fact, it has been found that a *continuous* transformation of heat into mechanical energy is possible only under the condition that at the same time heat shall be transferred by the working substance passing through the engine which performs the work, from a higher to a lower temperature, or, in other words, that the engine shall take in a certain amount of heat at a high temperature and give out a smaller amount of heat at a lower temperature.

188. Carnot's Cycle. A process, first described by Carnot, will illustrate the statement of the last article. It is a cycle consisting of four parts.

1. A gas, kept in constant contact with a reservoir at temperature $t_1°$, is expanded isothermally from P_1V_1 to P_2V_2, as represented (Fig. 97) by the curve *AB*. It absorbs an amount of heat H_1. External work, equal to the area *ABba*, is done by the gas.

2. The gas expands adiabatically, until its temperature has fallen to t_2. No heat is added during this step. The external work done by the gas is represented by area *BCcb*.

3. The gas is compressed isothermally, while in constant contact with a cool reservoir at temperature $t_2°$. An amount

of heat H_2 is given out, and the work done upon the gas is equal to the area $CDdc$.

4. The gas is compressed adiabatically until it again reaches its original condition, being heated to t_1. No heat is given out and the external work done upon the gas is represented by the area $DAad$.

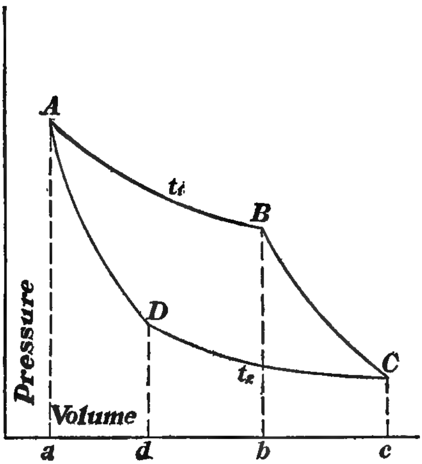


FIG. 97.

During the whole cycle external work, equal to the area $ABCD = ABba + BCcb - CDdc - DAad$, is done by the gas. This work has been obtained by a transformation of a part of the heat H_1 entering the engine. Therefore

$$H_1 - H_2 = W \quad (288)$$

The *efficiency* of the cycle is the ratio of the useful work to the total energy H_1 put into the machine. For this "theoretical" engine the efficiency may be shown to be

$$\frac{W}{H_1} = \frac{H_1 - H_2}{H_1} = \frac{T_1 - T_2}{T_1} \quad (289)$$

where T_1 is the absolute temperature of the hotter reservoir.

189. Irreversible Processes. While the work done by the engine might afterwards be used to restore some heat to the hotter reservoir, as, for example, by working the above cycle backwards, still there are some common processes in nature in which heat passes from higher to lower temperature without doing work. Such a process cannot be reversed. The most important example of this kind is the conduction of heat, as through the walls of the cylinder of a steam engine. After the heat has once reached the lowest temperature of all the surrounding bodies, it is impossible to obtain any further useful work from it.

Since conduction of heat cannot be avoided in steam engines, and since the heat thus transferred is lost, so far as useful work is concerned, it is evident that the efficiency of our actual steam engines must be smaller than that of the Carnot cycle, where no such loss was assumed to occur.

190. The Reciprocating Steam Engine. The first successful attempt to utilize the energy of steam was due to James Watt (1736–1819), whose apparatus took the form of the reciprocating engine. Fig. 98 illustrates the action of such an

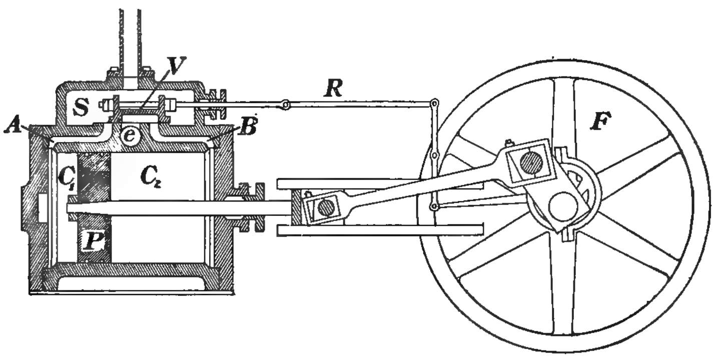


FIG. 98.

engine. Its essential parts are (*a*) the cylinder C_1C_2, in which a piston P moves back and forth; (*b*) a steam chest S, from which the steam, at high temperature, enters the cylinder, through openings A and B, called the ports; (*c*) the exhaust e, through which the steam, at a lower temperature, leaves, escaping either into the air or into a vacuum chamber, called the condenser.

A slide valve V is moved back and forth by the eccentric rod R, connected with the flywheel F, or some other rotating part of the engine. The motion of this valve is such that it connects the port on one side of the cylinder with the steam chest, and a short time afterwards the other port with the exhaust. With the position of the parts of the engine as shown in the figure, the pressure of the steam, entering from

the boiler, moves the piston to the right and sets the flywheel in motion. Shortly before the piston reaches its extreme position, the slide valve has moved far enough to the left to close both ports; the steam at the right-hand side is now compressed and its temperature rises. But in the meanwhile the slide valve has moved so far to the left that the steam will now enter through port B while the left cylinder is connected with the exhaust. Now the whole process is reversed, the piston returns to its original position and the cycle is repeated. The reciprocating action of the piston keeps up the motion of the rotating part of the engine

In some engines the process of expansion from the high pressure of the boiler to that of the exhaust is distributed over a number of steps, each successive cylinder being larger in diameter than the preceding one. Such engines are called compound, triple-expansion or quadruple-expansion engines according as the number of steps is two, three or four.

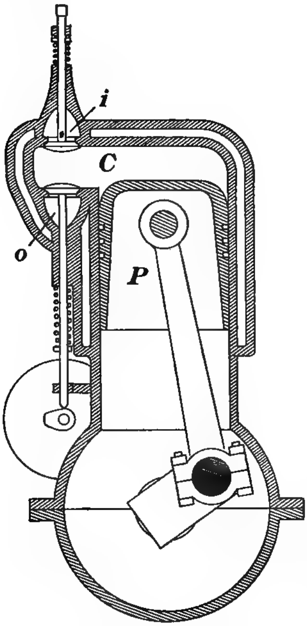


FIG. 99.

***191. The Internal Combustion Engine.** The best-known engines of this type are the gas and gasoline engines. Instead of leading steam from a separate boiler into the cylinder, a mixture of air and gas, or of air and gasoline which evaporates in the cylinder, is passed into the cylinder and there exploded. This produces the high temperature and the pressure necessary to push the piston forward. The cylinder is supplied with two valves, the inlet and the outlet valves, i and o (Fig. 99).

Both valves are closed when the explosion takes place producing the expansion stroke (1), which drives the piston forward. On its return (2), the outlet valve opens and the burnt

gases are ejected, after which this valve closes. During the suction stroke (3), the inlet valve opens and the fuel enters the cylinder. With both valves closed the piston returns, making a compression stroke (4), and compresses the gas mixture. Then the whole cycle repeats itself.

It is seen that four strokes or two complete to-and-fro motions of the piston follow each explosion. In its simplest form such an engine requires but a single cylinder. However, in order to increase the available power as well as to minimize the jar, due to the explosions, such engines are now furnished with two, three, four, and even six cylinders, all geared to the same shaft and delivering the thrusts, due to their individual explosions, at symmetrically periodic intervals. The action of such engines is characterized by remarkable speed and smoothness.

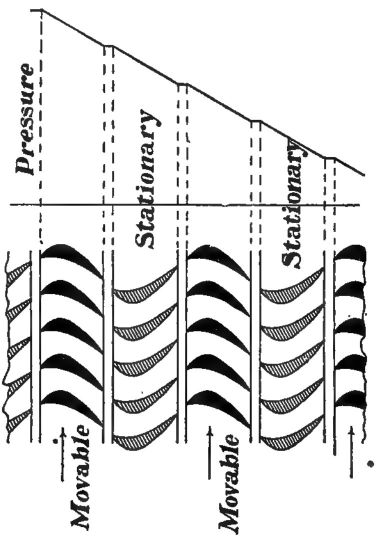


FIG. 100.

***192. The Steam Turbine.** The steam turbine consists of a revolving drum on whose periphery a large number of vanes are mounted (Fig. 100). Jets of steam are directed against these blades and by their impulses produce the rotation of the wheel. In the simplest types the transformation of heat into kinetic energy of the steam takes place at the nozzles through which the steam enters the turbine and the kinetic energy of the steam produces the useful work.

In the "multistage" turbine the expansion is divided into steps, the blades being further apart in successive stages. In these engines the steam, which is still under considerable pressure when striking the first set of vanes, also does work by its expansion as it passes through the turbine, and by this method

produces work inside the engine at the expense of the heat which it contains.

The angular velocity of steam turbines is very high, although it is lower in the multistage type than in the simple turbine. These machines occupy less space and run more quietly than the reciprocating engines and are especially adapted for driving alternating current dynamos and centrifugal pumps. On account of the absence of jarring they are coming more and more into use on board steamships.

Problems

1. What is the thermal capacity of 1 cu ft of air, expressed in British thermal units, the specific heat of air being 0.24 and the density 0.08 lb per cubic foot? *Ans.* 0.0192 B.T.U. per degree F.

2. A piece of copper, weighing 300 g and heated to 99°.4 C, is plunged into 400 g of water contained in a copper calorimeter whose mass is 90 g. The temperature of the calorimeter and its contents is raised from 20° to 25°.1 C. Find the specific heat of copper and the thermal capacity of the piece of copper introduced into the calorimeter.

Ans. $s = 0.0935$; $C = 28.04$ calories per degree.

3. To find the temperature of a certain furnace, a piece of platinum of mass 10 g is placed in it. After taking the temperature of the furnace it is suddenly plunged into 40 g of water at 10° C. The temperature of the water rises to 24° C. What is the temperature of the furnace, assuming the specific heat of platinum to be 0.032? *Ans.* 1774° C.

4. Calculate the thermal capacity of unit volume of mercury and of glass, the density of the latter being 2.5 g/cm³. Show that the thermal capacity of the immersed part of the thermometer may be taken without an appreciable error as numerically equal to 0.47 of its volume.

Ans. For mercury 0.449; for glass 0.50 calories per cm³ per degree.

5. One gram of anthracite coal if burned produces 7800 calories. How much heat, expressed in British thermal units, is produced by the burning of 1 lb of coal? *Ans.* 14,040 B.T.U.

6. How much heat is necessary to heat the air in a certain room, $6 \times 5 \times 3$ m from 0° C to 25° C? Assuming the mass of air to remain constant, how much water, cooling from 100° to 25° C, would furnish the heat required? *Ans.* (*a*) 698,220 calories.
(*b*) 9.31 kilos.

7. If it require half a horse power for 2 min to drill through a block of iron of 800 g mass, how much heat is produced? Supposing nine tenths of the heat to appear in the iron, how much does the temperature of the block rise? *Ans.* (*a*) 10,692.5 calories.
(*b*) 103°.7 C.

8. The falls of Niagara are 160 ft high. How much warmer should the water be at the bottom than at the top? *Ans.* 0°.206 F.

9. What must be the speed of a lead bullet if, upon striking a target, its temperature be raised from 27° C to its melting point, 327° C? Assume that all the heat produced serves to heat the bullet, and that the specific heat of lead is 0.031. *Ans.* 279 meters per second.

10. When a street car weighing 4000 kg and having a speed of 20 km per hour is stopped by the brakes, how much heat is produced? *Ans.* 14,750 calories.

11. Suppose the earth's rotation around its axis to be suddenly stopped. What change in temperature would be produced? Consider the earth a sphere of uniform density, and of specific heat 0.2; radius of earth, 6360 km. The moment of inertia of a sphere around any diameter is $2/5\,MR^2$, where M is its mass and R the radius. *Ans.* 51°.1 C. (If T be taken as 86,164 sec, *i.e.* one sidereal day, then $t = 51°.38$ C.)

12. If the average pressure in the cylinder of a steam engine be 10 kilograms-weight per square centimeter, and the area of the piston be 300 cm², how many calories does the steam lose when it pushes the piston 50 cm forward? *Ans.* 3511.7 calories.

13. How much heat is needed to produce the work done by a 100 H.P. engine running for one hour? *Ans.* 64.157×10^6 calories.

14. The efficiency of a condensing engine is about 16 per cent. How much coal is consumed by a 20,000 H.P. condensing engine in one hour, assuming 30 per cent of the heat of combustion to be lost in passing from the coal to the engine? (See problem 5.) *Ans.* 14,688 kilos.

15. The average locomotive has an efficiency of about 6 per cent. What horse power does it develop when consuming 1 ton of coal per hour? (See problem 5, assuming 1 ton to be 1000 kilos.) *Ans.* 729 H. P.

16. A perfect engine takes steam from a boiler at 150° C, and exhausts into a condenser at 30° C. Compute its efficiency. *Ans.* 28.37 per cent.

17. The mixture of air and gas in a gas engine reaches a temperature of about 1100° C when it is ignited, and the temperature of this mixture is reduced to 600° C by expansion. What would be the efficiency of a perfect engine working between these temperatures? *Ans.* 36.42 per cent.

18. If a compound marine engine consume 2 lb of coal per horse power every hour, what per cent of the energy of the coal is being transformed into work in the cylinder? *Ans.* 9.06 per cent.

19. Calculate from the specific heats of air, and, using equation (281), the value of the mechanical equivalent of heat. Density of air at 0° C = 0.001293 g per cm³. This calculation was first made by Mayer in 1842. *Ans.* $J = 4.1 \times 10^7$ ergs per calorie

CHANGE OF STATE

CHAPTER XXIV

FUSION

193. The Melting Point. It has already been shown that when the temperature of a solid reaches a certain point the body begins to melt, and that any further addition of heat, if not too rapidly applied, simply serves to hasten the melting process, without changing the temperature. If the process be interrupted by preventing heat from reaching or leaving the mixture formed, the temperature will remain constant and no further change in the relative amounts of solid and liquid takes place, thus showing that in this condition stable equilibrium exists.

The *melting point* or the *fusing point* of a substance is therefore *that temperature at which the solid and liquid states are in equilibrium under the existing pressure.* The fusing point is usually referred to atmospheric pressure. Above this temperature the substance exists as a liquid, while below this temperature it is usually a solid. When the change of state occurs at a relatively low temperature, the substance in question is generally known in the liquid state and we consequently speak of the temperature of transition as the freezing point. From the definition it is clear that *the freezing point and the melting point are one and the same temperature.*

Only crystalline bodies have definite melting points. Amorphous substances, such as glass, paraffine or wax, grow more and more plastic as the temperature is raised, and finally become liquid, hence no definite temperature can be found at which a transition from the distinctly solid state to the distinctly liquid state occurs. On account of this gradual change in plasticity such substances may be heated to softness and may then be

molded into any desired shape, or two pieces may even be welded together. By means of polarized light it may be shown that glasses soften sufficiently to permit of some molecular motion, and the removal of internal strains, at temperatures from 250° to 300° below the point at which the same glasses become fluid.

TABLE X

MELTING POINTS UNDER ATMOSPHERIC PRESSURE

SUBSTANCE	MELTING POINT	SUBSTANCE	MELTING POINT
Mercury	−38°.8	Aluminium	658°
Phosphorus	44°.3	Silver	960°
Sulphur	115°	Gold	1063°
Tin	232°	Copper	1083°
Bismuth	260°	Iron	1100°
Cadmium	320°	Steel	1350°
Lead	327°	Platinum	1755°
Zinc	418°	Iridium	2200°

194. Heat of Fusion. The quantity of heat necessary to melt a body of mass M is proportional to its mass, consequently we have

$$H = LM\ . \tag{290}$$

The proportionality factor

$$L = H/M \tag{291}$$

is called *the heat of fusion of the substance* and may be defined as the *heat per unit mass needed to change the substance from the solid to the liquid state, without change of temperature.* The heat of fusion is therefore a measure of the energy needed to produce this change of state. It is numerically equal to the heat absorbed in the fusion of one gram of the substance. It may easily be found by the method of mixtures. For water the value of L is nearly 80 calories per gram.[1]

The same amount of heat, LM calories, is liberated when a mass M *of the same liquid freezes.* Pails filled with water

[1] *For determination of the heat of fusion of water, see Manual, Exercise 43; for heat of fusion of tin, Exercise 45.*

are often placed in cellars in order that the water in freezing may liberate enough heat to prevent the freezing of the fruit or vegetables in the cellar. Ice keeps a refrigerator cool because it absorbs from the refrigerator and its contents the heat needed to melt the ice.

This absorption of heat during melting is employed in *freezing mixtures.* If salt be mixed with ice, a salt solution is produced. The heat of fusion of the ice must be supplied by the mixture and the vessels standing in it. Melting is therefore a cooling process, and freezing a warming process.

195. Supercooling. While a solid when heated always begins to melt as soon as the melting point is reached, a liquid, if carefully protected from mechanical disturbances, may frequently be cooled below this point without freezing. Thus water may easily be cooled to $-10°$, or more, without the formation of ice. This phenomenon is called supercooling. The liquid is then in a state of equilibrium less stable than that of the mixture of solid and liquid which will ensue as soon as a crystal is brought into the liquid.

This is easily shown with hyposulphite of soda. The salt, after having been melted at a temperature somewhat above 50°, will remain liquid in a stoppered flask for an indefinite time at ordinary temperatures. But the moment a crystal is dropped in, fine needles will be seen to shoot from it in all directions and soon the whole volume is filled with crystals. At the same time the temperature rises to the normal melting point, 49°.5. This evolution of heat is due to the heat of fusion liberated during the crystalization.

With some substances supercooling may be continued to such relatively low temperatures that the liquid becomes more and more viscous and finally solid, so far as its mechanical properties are concerned. Ordinary glass is an example of such a substance. Quartz crystals melt at about 1500°. If the molten mass be cooled, it becomes an amorphous, transparent substance, of valuable physical properties (Art. 155). Quartz glass, or fused quartz, is therefore nothing but supercooled liquid quartz.

196. Change of Volume during Fusion. In general, substances contract when they freeze and expand when they melt. Notable exceptions, however, are cast iron and type metal. While other bodies would, upon solidification, draw away from the mold, these metals expand, press into the mold, and reproduce the finest and most minute details.

The most common exception is water. One gram of water expands upon freezing from 1.00012 cm^3 to 1.092 cm^3. This expansion plays an important rôle in the disintegration of rocks during the winter season. The pressure exerted by freezing water is very great. A cast-iron bomb if filled with water and securely sealed by a screw cap may, when placed in a freezing mixture, explode with a loud report.

197. Influence of Pressure upon the Freezing Point. All substances which expand upon melting have their melting points

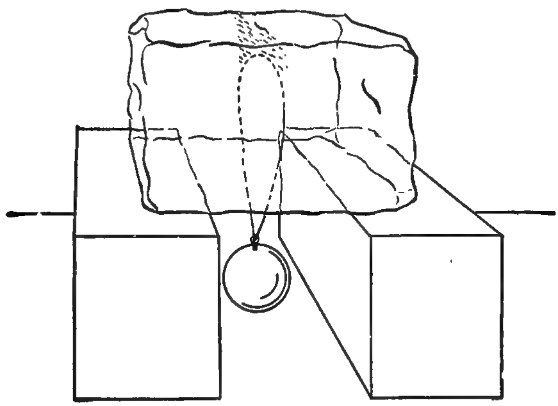


FIG. 101.

raised by an increase of pressure; all substances which contract upon melting have their melting points lowered. Or, in general, *pressure favors that state in which the volume is least.* This influence of pressure is very small. Thus an increase of pressure of one atmosphere lowers the melting point of ice only 0°.0075 C.

Snow easily packs under pressure if the temperature be near the freezing point, but will not do so if the weather be too cold. Two pieces of ice may be frozen together under warm water by

applying considerable pressure and then releasing them suddenly.

Let a mass of several kilos be hung by a thin wire over a block of ice (Fig. 101). In a short time the wire will be found to have cut completely through the block, leaving the ice as solid as it was at the beginning. The ice just below the wire melts on account of the increased pressure and it absorbs heat of fusion by which a temperature a little below 0° is produced at the place where the wire touches the ice. The water passes the wire and freezes again above it, on being released from the pressure. If the mass be of metal, it will be better to insert a link of string between the mass and the fine wire.

198. Freezing Point of Solutions. A small quantity of salt dissolved in water lowers the freezing point nearly in proportion to the quantity dissolved. When such a solution begins to freeze, however, we find that it is only the *water* which freezes out, while the salt, remaining in solution, makes it more concentrated and consequently lowers the freezing point still further.

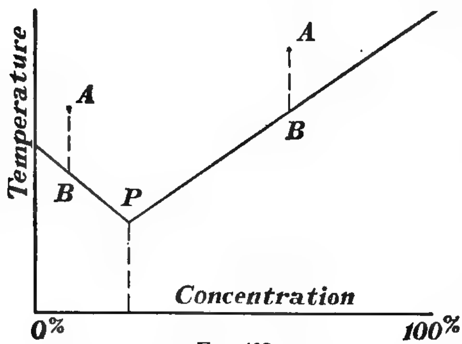


FIG. 102.

If, on the other hand, a concentrated salt solution be cooled, the *salt* crystallizes out as soon as the limit of its solubility is reached, and the solution becomes more dilute. The less concentrated the solution, the lower the temperature at which the salt begins to "freeze" out.

In the first case the solution becomes saturated with respect to the solvent, water; in the second, with respect to the dissolved substance, salt. If the temperature at which saturation occurs be plotted as a function of the concentration of the solution, two curves (Fig. 102) are obtained which meet at a point P of minimum temperature. Cool a solution of any concen-

tration, for example, one corresponding to a point A. Crystallization, either of the salt or of the solvent, will begin as soon as the temperature has fallen to that of some point B on the saturation curve. Upon further cooling, the solution moves along the saturation curve to the point P, where the whole mass solidifies as a *mechanical* mixture of ice and salt. Such a mixture is called a *cryohydrate*. The cryohydrate of common salt and water contains 26 per cent of salt and solidifies at $-23°$ C.

Alloys of metals behave in a manner similar to solutions. In general, the melting point of the solvent is lowered by the addition of a small quantity of some other metal. Solders are well-known examples of this fact.

A notable example of this is found in Wood's metal, composed of 52 per cent of bismuth, 26 per cent of lead, 15 per cent of tin, and 7 per cent of cadmium. This alloy melts at 75°.5 C, so that a spoon made of this metal melts when placed in hot tea.

CHAPTER XXV

VAPORIZATION

199. Vaporization. Vaporization is the process of transforming a substance from the solid or the liquid state into the gaseous state. According to the molecular theory some of the molecules of a solid or of a liquid possess sufficient kinetic energy to carry them through and beyond the surface of the mass. The molecules thus "freed" form the vapor filling the surrounding space. By virtue of their impact upon the restraining walls they exert a definite pressure. This pressure is termed the vapor pressure of the substance for that temperature.

Vaporization from a liquid is called *evaporation.* Vaporization from a solid is termed *sublimation.*

200. Evaporation. If a small amount of liquid be introduced into a barometric tube above the mercury, it begins to evaporate and, on account of the pressure which the vapor exerts, the mercury falls to a certain height which is independent of the width of the tube and of the space filled with vapor, so long as there is any liquid left in the tube (Fig. 103). Vapor which is in equilibrium with its liquid is called *saturated vapor.* The molecular theory teaches that after equilibrium is established, as many molecules leave the liquid during a given time as reënter it from the vapor. If the space above the liquid be increased, more vapor is formed, but so long as any liquid is present, the meniscus of the mercury in the tube remains at the same height above the level of the mercury in the cup. This shows that *satu-*

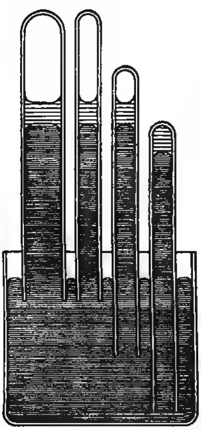


Fig. 103.

rated vapor at a given temperature always exerts the same pressure.

If the volume be increased until all liquid is evaporated, the vapor becomes isolated and the pressure at a given temperature depends only upon the volume occupied by the vapor. It is then called an unsaturated or superheated vapor and follows very nearly the gas law.

In order to distinguish the pressure exerted by an unsaturated vapor from that of a saturated vapor we shall call the pressure due to the latter *vapor tension.*

201. Evaporation and Dalton's Law. When water is introduced into a barometric tube at a temperature of 20° C, the mercury meniscus falls 1.74 cm. The vapor tension of water at 20° corresponds, therefore, to 1.74 cm of mercury, or it is equal to 23,170 dynes per square centimeter. If the tube had been partially filled with some gas which does not act chemically upon water, the lowering of the mercury upon the introduction of water would have been very nearly the same as before. This shows that the vapor tension of the water has simply been added to the pressure of the gas previously in the tube. This illustrates Dalton's law (Art. 95).

A liquid evaporates into a vacuum or into any mixture of gases until the individual pressure produced by its vapor equals the vapor tension of the liquid at the existing temperature. The only influence of the other gases will be a decrease in the rate of evaporation. Thus water at 20° C evaporates so long as the partial pressure of water vapor in the surrounding atmosphere is below 1.74 cm of mercury, or until the atmosphere becomes saturated with vapor.

202. The Vapor Tension Curve.[1] Change in temperature has a great influence upon the pressure exerted by a saturated vapor. The relation between the two may be plotted as a curve on a pressure-temperature diagram (Fig. 104). The points on this curve will then represent the condition of equilibrium between the liquid and its saturated vapor. Such a

[1] *For determination of the vapor tension of ether, see Manual, Exercise 46.*

curve is called the *vapor tension* or *evaporation curve*. If the pressure of the vapor in contact with its liquid be greater than the vapor tension of the liquid at the same temperature, the vapor will condense. If the pressure be smaller, the liquid will evaporate. For conditions of equilibrium, therefore, the space on the left-hand side of the curve represents the liquid state alone, while that on the right represents the gaseous state. Fig. 104 represents these relations for water.

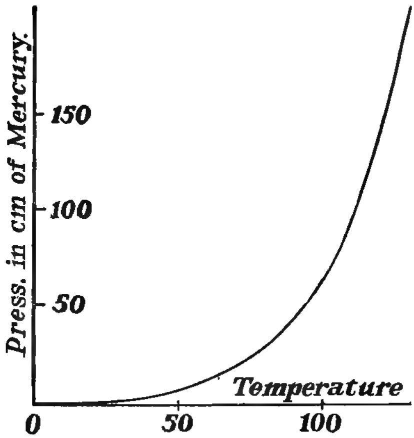


FIG. 104.

203. The Boiling Point. When a liquid is heated, its temperature rises and its vapor tension accordingly increases, until finally bubbles are formed in the liquid itself, especially at the place where the heat is applied. The liquid "boils" when evaporation no longer takes place quietly on the surface. The bubbles of saturated vapor expand against the pressure of the surrounding atmosphere and that of the small layer of liquid above them as they rise to the surface, while at the same time the liquid rapidly evaporates *into* the bubble from all sides.

The boiling point of a substance is the temperature at which the vapor tension of the substance equals the gas pressure upon the liquid, no matter to what this pressure may be due. The "normal" boiling point is always referred to a pressure corresponding to 760 mm of mercury.

The boiling point is evidently the temperature on the vapor tension curve corresponding to the gaseous pressure on the liquid. It is given accurately by the reading of a thermometer hung in the vapor a short distance above the boiling liquid.[1]

[1] *For method of determination of the boiling point, see Manual, Exercise 37.*

TABLE XI

BOILING POINTS OF SOME LIQUIDS UNDER ATMOSPHERIC PRESSURE

SUBSTANCE	BOILING POINT	SUBSTANCE	BOILING POINT
Ethylene	− 103°	Chloroform	+ 61°
Ammonia	− 38°.5	Alcohol	+ 78°
Chlorine	− 33°.6	Toluene	+ 110°
Ether	+ 35°	Glycerine	+ 290°
Carbon bisulphide . .	+ 46°	Mercury	+ 357°

204. Superheating. The vapor bubbles form in a boiling liquid usually at places where minute air bubbles adhere to the walls of the vessel or to some foreign substances present in the liquid. After the air has been removed by previous boiling, the liquid may often be heated considerably above the boiling point, since no opportunity is given for the formation of vapor inside the liquid. The liquid is then said to be superheated. In such cases sudden boiling will finally set in with almost explosive violence.

In order to show the superheating of a liquid, heat a beaker full of water that has been boiled for a few minutes and allowed to cool. The temperature may be carried several degrees above the boiling point corresponding to the reading of the barometer. If now there be added to the water a small quantity of white sand or finely powdered glass, violent boiling will ensue.

As another illustration of superheating, fill a tube closed at one end and about 80 cm long, with mercury and a few cubic centimeters of ether. Invert it in a deep dish filled with mercury. If care be taken to remove all air, the ether will remain liquid on the top of the mercury. The experiment succeeds often with a mercury column more than 76 cm long, the liquid being actually under a pull and yet not evaporating. A slight jar, however, will start evaporation, and the mercury will rapidly fall to a position corresponding to the normal pressure given by the vapor tension curve ; that is, 44 cm at 20° C.

A superheated liquid is therefore in a less stable state than the mixture of the liquid and its vapor at the same temperature.

205. Influence of Pressure upon the Boiling Point. If a vessel with water at ordinary temperature be placed under the receiver of an air pump, a few strokes of the pump will cause the water to boil. The experimental result that the boiling point is lowered by a decrease of pressure could have been predicted from a study of the general shape of the vapor tension curve and from the definition of the boiling point as given (Art. 203).

The influence of pressure upon the boiling point is much more marked than its influence upon the melting point. The boiling point of water, for example, changes 0°.37 for a change of 1 cm of mercury in barometric pressure.[1]

Since the change of barometric pressure amounts to about 1 mm of mercury for a vertical rise of 11 m, water boils on Pike's Peak, 4310 m above sea level, at 85°, a temperature at which many ordinary cooking operations are impossible. On the other hand, the boiling point of water in a steam boiler under a pressure of 100 lb per square inch is 155° C.

TABLE XII

BOILING POINT OF WATER UNDER DIFFERENT BAROMETRIC PRESSURES

Pressure	Boiling Point	Pressure	Boiling Point	Pressure	Boiling Point
73.0	98°.88	75.0	99°.63	77.0	100°.37
73.5	99°.07	75.5	99°.82	77.5	100°.55
74.0	99°.26	76.0	100°.00	78.0	100°.73
74.5	99°.44	76.5	100°.18	78.5	100°.91

206. Vapor Tension and Boiling Point of Solutions. The vapor tension of a solution containing a non-volatile salt is always lower than that of the solvent at the same temperature. For dilute solutions the decrease is proportional to the concentration. The vapor tension curve of a solution, therefore, always

[1] *For effect of pressure upon the boiling point, see Manual, Exercise 47.*

lies to the right of that of the solvent. In Fig. 105 the vapor tension of a solution is indicated by the dotted lines, while the full line gives the vapor tension of the solvent. Consequently for a given pressure, the boiling point *M* of the solution is higher than *N*, that of the solvent. For example, 35.5 parts of sodium chloride (common salt) dissolved in 65.5 parts of water lowers the vapor tension at 100° nearly 18 cm of mercury and raises the boiling point to 107°.5.

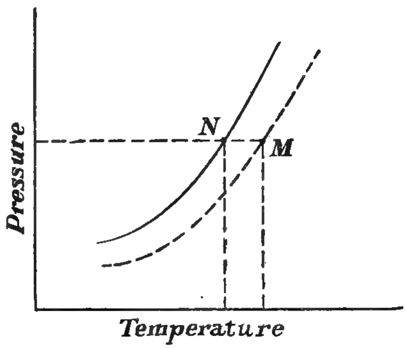


FIG. 105

It is frequently stated that the temperature of the vapor above the boiling solution is the boiling point of the solvent. A thermometer hung in the vapor condenses some of the vapor, and then indicates the boiling point of the solvent. But if such condensation be prevented, it may be shown that the vapor of a boiling solution is at a temperature equal to the boiling point of the solution and not to that of the solvent.

When a solution, containing non-volatile substances is boiled, only the solvent evaporates, while the dissolved substance crystallizes out, when the solution becomes sufficiently saturated.

207. Distillation. The vapor rising from a boiling liquid may be condensed into a liquid by being passed into a cold vessel. This combination of boiling and condensation is called distillation. It affords a convenient method for freeing a liquid from impurities, as water from salts, or mercury from other metals. Evaporation in a vacuum is employed either when the normal boiling point is very high as in the case of mercury, or when the substance crystallizing out is chemically changed at the normal boiling point. Sugar solutions are evaporated in a vacuum to prevent scorching the sugar at the boiling point of the syrup.

The saturated vapor above a mixture of two volatile

liquids, such as alcohol and water, contains in general the two constituents in a different proportion from that in the solution. If such a solution be boiled, the constituent whose percentage is higher in the vapor than in the liquid, distills over more rapidly and consequently is found in much larger concentration in the distillate than in the original solution.[1]

In some mixtures there exists a definite concentration at which the proportion of the constituents is the same in both, liquid and vapor. Solutions of this concentration distil over unchanged. An example of this is common alcohol of 96 per cent concentration, which boils at 78°.17 C, a temperature 0°.13 C below the boiling point of pure alcohol. At higher concentration than 96 per cent, there is relatively more water in the vapor than in the liquid, and, consequently, more water distills over than alcohol.

It is evident that alcohol containing much water may be concentrated by repeated distillation up to a strength of 96 per cent, but that it is impossible to obtain by this method an alcohol of greater concentration. Absolute alcohol is obtained by allowing 96 per cent alcohol to stand for some time over quicklime.

208. Heat of Vaporization.[2] From the foregoing considerations it is clear that when a liquid is changed to a vapor, a certain quantity of heat is needed to effect this transformation. *The heat of vaporization of a substance denotes the heat per unit mass needed to vaporize that substance without change of temperature.* This heat of vaporization is constant for a given substance for a given temperature, but *decreases* as the temperature *increases.* For water this relation is given by the equation, proposed by Griffiths,

$$L = 596.6 - 0.601\,t \tag{292}$$

[1] *For a disscusion of the behavior of different mixtures of two volatile liquids, see Chwolson, Lehrbuch der Physik, vol. III, p. 934.*

[2] *For determination of the heat of vaporization of water, see Manual, Exercise 44.*

More recent determinations are found in the following table:[1]

TABLE XIII

HEAT OF VAPORIZATION OF WATER

TEMPERATURE	L IN CAL. PER GRAM
0° C	596.3
25° C	582.5
50° C	568.2
75° C	553.3
100° C	538.0

Since water, on vaporization at 100°, expands to about 1650 times its liquid volume, it follows that by this expansion it has done work and has absorbed energy. Of the total heat energy absorbed in the vaporization of water, about $\frac{1}{13}$ is needed to effect this expansion, while the remaining $\frac{12}{13}$ is to be regarded as an increase in the potential energy of the water molecules, so long as they exist in the form of steam. When the steam condenses again to water, all the heat absorbed during vaporization is given out. Upon this principle depend many important industrial applications, such as steam heating, steam cooking, etc.

209. Cooling by Evaporation. Owing to the large amount of heat absorbed, rapid evaporation is a very efficient means of cooling. To that end we sprinkle the floors or sidewalks on a hot day, or bathe with water or alcohol patients suffering from fevers. Water kept in a porous jar is always at a lower temperature than that of the surrounding air, owing to the rapid evaporation of part of the water on the outer surface of the porous vessel.

Water may even be frozen by its own evaporation provided this be made sufficiently rapid. To this end a small flat cup of water, thermally insulated, is placed over a shallow dish containing concentrated sulphuric acid (Fig. 106), and the whole covered by a flat receiver upon the plate of an air pump. At

[1] A. W. Smith, *Monthly Weather Review*, October, 1907.

the first stroke of the pump a cloud of mist is seen which is largely absorbed by the acid. After the air has been pumped out of the water, the pressure in the receiver is soon reduced to a value below the vapor tension of water at room temperature, and rapid boiling begins. Now, since all heat needed for the vaporization must be supplied from the cup and its contents, it is evident that the temperature of the water must fall quickly, and if the pump be worked rapidly, the water boils and freezes at the same time. For success in this experiment the pump must not only exhaust rapidly, but must also reduce the pressure in the receiver to something less than 4.6 mm of mercury (Art. 211).

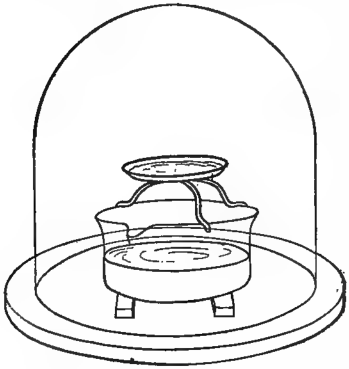


FIG. 106.

210. Cooling by Expansion of Gases. We have already seen (Art. 175) that if a gas be heated under constant pressure, it expands, and absorbs heat. The converse of this truth is seen in the fact that if a gas under pressure be allowed to expand it tends to absorb heat, and its temperature rapidly falls. If the gas be liquefied and held under great pressure, then on its release we have the combined cooling effects due to vaporization of the liquid and the expansion of the resultant vapor. In this way a gas may be cooled so suddenly as to be frozen solid.

If a cylinder containing liquid carbon dioxide be placed in a vertical position with the valve down, then on opening the valve, the liquid will be driven out by the pressure of the confined gas. Owing to the high vapor tension of the liquid, a very rapid evaporation and expansion occurs, and the gas is quickly chilled to the freezing point. If a bag made of flannel or chamois skin be held over the opened valve, it will soon be filled with a snowy substance, the solid carbon dioxide, which under atmospheric pressure has a temperature of $-78°$. If this snow be gathered up and compressed into a brick, it may be kept

for hours in the open air. It slowly sublimes without passing through the liquid state (Art. 211).

A mixture of solid carbon dioxide and ether or alcohol is more convenient for experimentation than the solid dioxide alone, because the liquid at $-78°$ insures better thermal contact with bodies immersed in it. With this mixture mercury may readily be frozen.

In commercial refrigerating plants liquid ammonia is rapidly evaporated in a system of coils by the action of a pump which constantly draws off the ammonia vapor, and at the same time compresses it at another part of the apparatus, the condenser, which is cooled by running water. This takes up the heat of condensation. Under the high pressure in the condenser the ammonia liquefies and then by a regulating valve is slowly readmitted to the coils in which the evaporation takes place. This part is called the evaporator and is usually immersed in a tank filled with brine. The brine is kept in circulation by separate machinery. In artificial ice plants it flows around the vessels containing the water to be frozen; in the cold storage plants it is pumped through coils placed in the rooms which are intended to be kept cool.

211. Sublimation. The pressure of the vapor produced by a solid is usually quite small. For most solids it is practically zero at ordinary temperatures, although our sense of smell often tells us that some vapor is being given off. For ice at 0° the saturated vapor pressure is 4.6 mm of mercury, and decreases rapidly with decrease of temperature.

For some solids the saturated vapor pressure may become quite large at higher temperatures. For iodine it is only 0.01 mm of mercury at 0°, but 47.5 mm at 100°, and 687.2 mm, or almost one atmosphere, at 180°.

As in Art. 202 we may plot a saturated vapor pressure curve representing the conditions under which a solid is in equilibrium with its vapor. We shall call this curve the sublimation curve.

212. The Triple Point. We have seen in Arts. 197, 202, and 211, that under certain conditions of pressure and temperature

two states of a substance may be in equilibrium, and that these conditions are represented on a pressure-temperature diagram by curves; namely, the fusion curve, the vapor-tension curve and the sublimation curve. In Fig. 107 all three of these curves are drawn for a substance whose melting point decreases with increase of pressure. The curves intersect at a point which is called the triple point and indicates the pressure and temperature at which *all three states are in equilibrium.* For water this point is at the temperature $+0^\circ.0076$ and at a pressure of 4.6 mm of mercury.

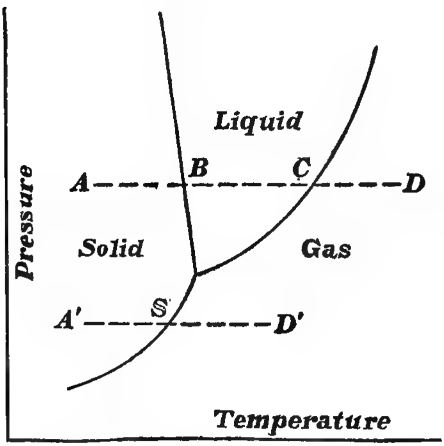


FIG. 107.

The temperature of a solid body when heated under a pressure higher than that of the triple point will first rise, line *AB* (Fig. 107), until the fusion curve is reached. Then it will melt while the temperature remains constant. After the melting is completed the temperature of the liquid will rise, line *BC*, until the vapor-tension curve is reached; but unless superheated, it cannot remain liquid beyond *C*, which is the boiling point for a given pressure. After all liquid is evaporated, further heating will increase the temperature of the vapor formed, line *CD*.

If the triple point of a substance lie at a pressure of more than one atmosphere, the solid, when heated in an open vessel, will not melt, but will sublime, as is shown in the figure by the line *A'SD'*. Camphor and carbon dioxide belong to this class of substances. They have no "normal" melting or boiling point, since under atmospheric pressure the liquid state is impossible for a state of equilibrium.

CHAPTER XXVI

HYGROMETRY

213. The Dew Point. The atmosphere is said to be saturated with moisture if the partial pressure of the water vapor in the air alone equals the vapor tension of water at the temperature of the air; that is, if the water vapor becomes saturated. In this case no further evaporation takes place from a wet surface. In general, the atmosphere is not saturated, but if the air be cooled, it may reach a temperature at which it is saturated. Cooling beyond this point results in a condensation of the vapor into water. The dew point is the temperature at which the water vapor present in the air becomes saturated.

The *dew point hygrometer* consists of a vessel with an outer surface of highly polished metal, and containing some volatile liquid, such as ether. By rapid evaporation of this liquid the vessel may be cooled below the dew point. This fact is easily recognized by the formation of a thin film of minute water drops upon the metallic surface. The temperature of the vessel is measured by an accurate thermometer. The vapor tension corresponding to the dew point gives at once the actual vapor pressure in the surrounding air.

Dew does not "fall," but is produced by the condensation of moisture upon surfaces which have been cooled below the dew point. This happens frequently on clear nights when the earth loses much heat by radiation to the sky. Dew forms on the upper surfaces of stones, plants, etc. Moisture is also frequently found collected under cold stones, because the soil underneath remains warm and the air in contact with it has a dew point above the temperature of the stone.

Frost is formed in the same manner as dew, except that the partial pressure of the water vapor is below that of the triple

point, so that the vapor will not condense until a temperature below 0° has been reached.

214. Relative Humidity. Relative humidity is the ratio of the pressure of water vapor in the air to the saturated vapor pressure at the temperature of the air. After the dew point has been determined it is only necessary to read from the vapor tension curve the pressure corresponding to the dew point and divide this by the vapor tension corresponding to the temperature of the air. This ratio is the *relative humidity.*

Our sense of dryness and dampness does not depend upon the absolute amount of water vapor in the air. At 0° air is saturated with moisture, when the partial pressure is 4.57 mm of mercury. At 0° air containing this amount of water will feel very moist. At 25° the partial pressure of the same amount of water in the air would be (eq. 258) nearly

$$P = \frac{4.57}{273} 298 = 5.0 \text{ mm}$$

of mercury; but the vapor tension at this temperature is 23.52 mm, so that in this case the relative humidity would be only $\frac{5.0}{23.52}$ or 21.3 per cent, which would make the air appear quite dry. The dew point would be found to be 1°.2. The capacity of the air for water vapor *nearly doubles for every rise in temperature of 10° C.*

The relative humidity may also be determined by the use of the wet and dry bulb thermometers, or the psychrometer. This instrument consists of two thermometers, the bulb of one being surrounded by a wet piece of muslin, or by a wick dipping into a vessel of water. The drier the air, the more rapidly does the water around the wet bulb evaporate and the lower will be the temperature of this thermometer. The evaporation on the wet bulb may be facilitated by whirling the thermometer through the air (sling thermometer). Tables have been prepared giving direct readings of the relative humidity, from the difference in the readings of the two thermometers and the temperature of the air as determined by the dry bulb thermometer.

*215. **Condensation of Water in the Atmosphere.** We have seen that when air is cooled below the dew point the water vapor begins to condense upon any solid or liquid surface. There are always numerous dust particles in the air which form the nuclei for the water drops. The greater the number of nuclei the smaller are the original droplets. Very small drops remain suspended in the air for a long time and form mists or fogs. These are frequently formed when moist layers of air are driven by the wind over cold water surfaces, or if they are cooled rather slowly. The dense fogs in large cities are explained by the large number of dust and soot particles in the air.

The condensation accompanying the cooling of moist air may easily be shown by the sudden expansion of moist air under the receiver of an air pump. If a light be viewed through the mist while it forms, beautiful color effects may be observed.

Clouds may be formed in two different ways:

(*a*) The "cumulus" clouds are due to the condensation of water vapor when moist air is carried by an upward air current to regions of lower pressure and its temperature is lowered by adiabatic expansion. These clouds appear in billowy, well-defined shapes, and are due to local disturbances in the atmosphere. They are rarely more than three miles high.

(*b*) The "stratus" clouds are probably due to the mingling of a cold air current with a warmer damp current, either overflowing it or driven up towards it, and to an additional cooling by expansion. The result is an extensive layer, without well-defined shape, covering the sky. These rain clouds seldom have an elevation of more than two miles. The "cirrus" clouds are formed at an elevation of about seven miles.

Very small water drops floating at high altitudes in the air produce the optical effect known as a "corona," while a "halo" ~~is~~ a similar effect caused by particles of ice.

The amount of rainfall in different parts of this country varies from 5 to 10 inches per year in the western part of Arizona, Nevada and Utah, to over 80 inches on the northern Pacific coast. The largest rainfall in the Atlantic states is 70 inches per year, in the mountains of North Carolina. The average rainfall in Michigan is between 25 and 35 inches per year.

CHAPTER XXVII

LIQUEFACTION OF GASES

216. Liquefaction by Pressure. We have seen (Art. 202) that condensation of vapor will in general occur at a given temperature when the pressure upon the vapor is raised to a value higher than that which corresponds to this temperature on the vapor tension curve. Vapors are gases, but the term *vapor* is commonly used, if the temperature be not far removed from the point of condensation.

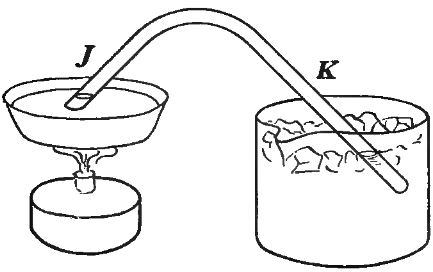


FIG. 108.

Some gases may be easily liquefied at ordinary temperatures by the application of pressure alone, as, for example, chlorine, ammonia and sulphur dioxide. Faraday[1] combined a cooling of the gas with compression, and thus liquefied carbon dioxide.

His apparatus (Fig. 108) consisted of a bent tube, into one end of which, *J*, he had sealed the chemicals for producing the gas, while the other end, *K*, was inserted in a freezing mixture. The gas, when generated, was thus liquefied under its own pressure. Other gases have for a long time withstood all attempts at liquefaction by simple methods.

217. The Critical Point. In 1863 Andrews discovered the fact that carbon dioxide cannot be liquefied by any pressure, however large, if its temperature exceed 31°.1. His experiments[2] may best be described by reference to a diagram (Fig. 109) in which the state of the substance is represented by

[1] Faraday, *Phil. Trans., 1823.* [2] Andrews, *Phil. Trans., 1869.*

its volume and the corresponding pressure. Starting with the carbon dioxide gas at a low temperature, say 13°.1, and compressing it isothermally, the pressure constantly increased until it equaled the vapor tension at 13°.1, corresponding to point *B*. At this point liquefaction began. The whole volume of the mixture of liquid *and* gas now decreased while the pressure remained constant, until all the gas was liquefied, point *C*. Upon further compression the pressure of the liquid rapidly increased.

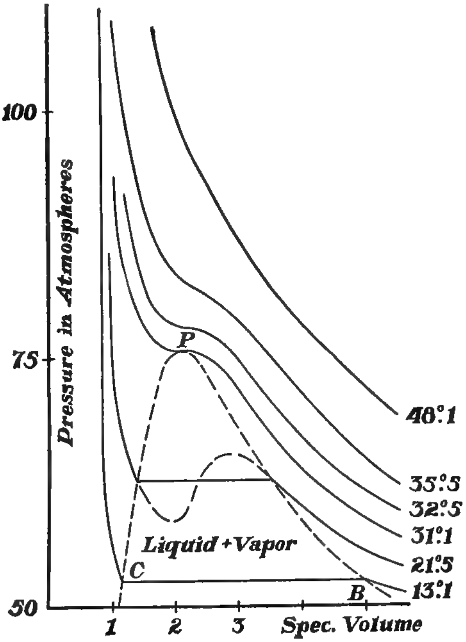


FIG. 109.

Similar curves were obtained at higher temperatures, but it was found that the higher the temperature, the shorter became the horizontal part of the curve *BC*, which represents the mixture of liquid and gas. At 31°.1 no visible liquefaction could be obtained and the curve showed only a distinct bending toward the horizontal before suddenly rising.

At still higher temperatures the curves became smooth and similar to those of a substance obeying the gas law. The area enclosed by the dotted line shows then the conditions under which a mixture of liquid and gas can exist. The temperature of the isothermal on which the substance upon compression ceases to show a free surface, which alone shows a distinction between the liquid and the gaseous state, is called *the critical temperature*. The point of contact *P* of the region of mixture with this isothermal line is called *the critical point;* and the

corresponding pressure and specific volume are termed the *critical pressure* and *critical volume* for the substance.

218. Transition through the Critical Point. Let a heavy walled glass tube (Fig. 110), closed at both ends, contain a definite amount of ether, the space above the liquid being occupied by the vapor. The ether is clearly in a state corresponding to some point in the region of mixture. If the tube be slowly heated, it will be found that the liquid expands first slowly, then more rapidly. This is accompanied by a considerable increase of pressure. The meniscus of the liquid becomes flatter, and, as the critical point is reached, it becomes indistinct and finally disappears. The tube is now to all appearances filled with a homogeneous substance, the ether having passed out from the region of mixture. It is idle to discuss the question whether it is now a liquid or a gas. At this state we cannot distinguish between the two.

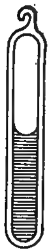


Fig. 110.

When the temperature is allowed to fall again, a hazy cloud forms in the upper part of the tube, the meniscus reappears, and the substance is again in two distinct states of aggregation, liquid and gaseous.

Table XIV

Critical Temperature and Pressure

Substance	Crit. Temp.	Crit. Pressure
Carbon dioxide	31°.1	77 atmos.
Ammonia	131°	113 "
Chlorine	146°	93.5 "
Ether	190°	37 "
Alcohol	240°	64 "
Water	374°	195 "

***219. Van der Waals's Equation.** It is apparent from the form of the curves (Fig. 109) that the gas law does not accurately represent the state of a gas near the region of liquefaction. In

1879 van der Waals proposed[1] an equation which, while not quite exact, yet represents the actual curves much better than the gas law, not only at the right-hand side of the region of mixture, but also on the left-hand side. In the region of mixture this equation gives a curved line as shown by the dotted line (Fig. 109) for the isothermal line corresponding to 21°.5.

Van der Waals's equation is

$$\left(p+\frac{a}{v^2}\right)(v-b)=RT \tag{293}$$

in which a and b are constants characteristic of the gas in question provided v represent the specific volume of the gas. This equation, with changed values for a and b, has been found to be applicable to many other gases besides carbon dioxide.

This equation becomes the gas law when the correction terms are small, that is, when the volume is very large in comparison with the volume just before condensation, or when the temperature is very high in comparison with the critical temperature.

The constant a was introduced in order to take into account the attraction between the molecules of the gas. This attraction will evidently produce a smaller volume, and its effect is equivalent to a pressure which must be added to the external pressure. The attraction between the molecules is proportional to the number of molecules, exerting the attraction, as well as to that of the attracted molecules, or it is proportional to the square of the total number of molecules present. But at constant temperature their number varies directly as the density of the gas and inversely as its volume. The correction factor which must be added to the external pressure is thus equal to $\frac{a}{v^2}$.

The fact that the molecules, however small, still have *some size*, and consequently occupy some small volume, is taken into account by subtracting a small volume, b, from the measured volume. It should be mentioned, however, that b is not the

[1] *Translated in Phys. Mem., Phys. Soc. London, vol. 1.*

actual volume of the molecules, but represents, as van der Waals has shown, a volume four times as large.

***220. The Regenerative Process.** Some gases possess a critical temperature below $-100°$ and must therefore be cooled be-

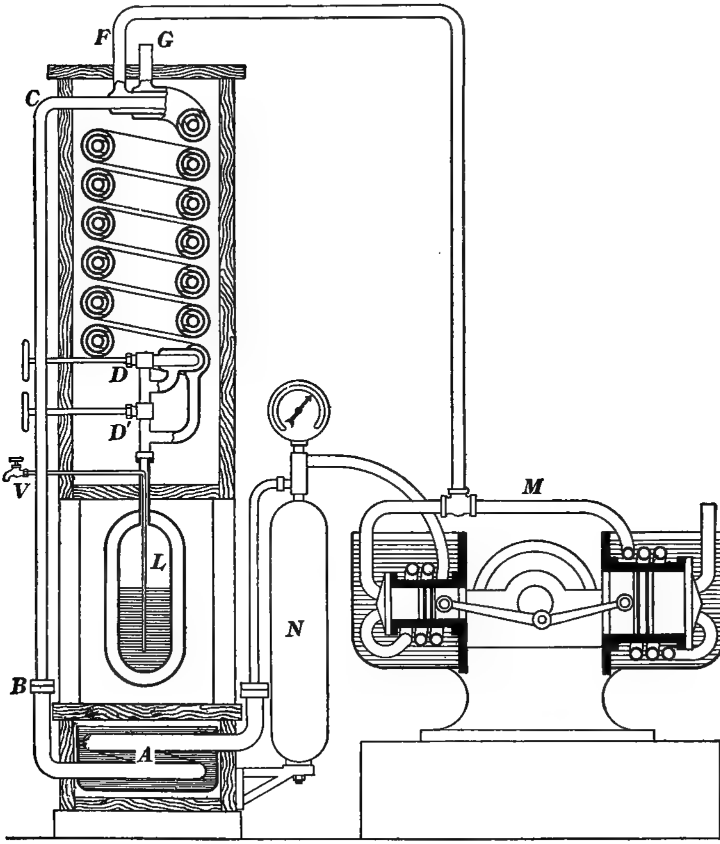


FIG. 111.

low this temperature before they can be liquefied. A method which allows a continuous production of these substances in the liquid state was discovered in 1895 by Linde and at about the same time by Hampson.

The principle used is based upon the fact that gases cool

upon expansion (Arts. 181 and 186). The gas is highly compressed to a pressure of 200 atmospheres in a coil immersed in a low temperature bath *A* (Fig. 111). It is then led through the pipe *BC* to the inner one of three concentric spiral tubes, and finally expands through a needle valve *D* to a lower pressure, 10 to 20 atmospheres. The cooled gas is then led back through the second one of the concentric spiral coils which completely surrounds the high pressure coil. The outflowing gas consequently cools the compressed gas before this gas reaches the expansion valve. Thus, as the process is continuous, the temperature of the compressed gas is constantly being lowered, until it finally reaches the critical temperature. The gas is then partly liquefied and the liquid collects in the vessel below the valve *D′*, through which the cold gas expands from about 20 atmospheres to atmospheric pressure. The non-liquefied portion of this gas below *D′* is led back through the outer one of the three tubes and thus assists in cooling the compressed gas still further.

Liquid air machines built on this principle may now be found in many physical laboratories. Hydrogen was first liquefied by Dewar in 1898, and frozen to a foamlike solid by boiling it under reduced pressure.[1] Helium was first liquefied by Kammerlingh-Onnes[2] in 1908. By boiling this liquid under reduced pressure, he attained the temperature – 270°, or *within 3° of the absolute zero.*

TABLE XV

FREEZING POINT, BOILING POINT AND CRITICAL DATA OF GASES

SUBSTANCE	FREEZING POINT	BOILING POINT	CRIT. TEMP.	CRIT. PRESSURE
Helium	——	– 268°.8	– 268°	2–3 atmos.
Hydrogen . . .	– 260°	– 252°.5	– 240°.8	14 "
Nitrogen . . .	– 210°	– 194°	– 145°	34 "
Oxygen	– 227°	– 181°	– 118°	50 "
Ethylene . . .	——	– 103°	+ 10°	38 "

[1] Dewar, *Chemical News, 1900.*   [2] Kammerlingh-Onnes, *Nature. 1908.*

Problems

1. Equal masses of water at 25° C and ice at 0° C are mixed. How much ice remains? *Ans.* 0.6875.

2. A mass of 100 g of melting ice is placed in a copper calorimeter whose mass is 100 g, and which contains 500 g of water at 30° C. After the ice is all melted, what will be the temperature of the calorimeter? *Ans.* 11°.95 C.

3. A mass of 100 g of ice at −10° C is dropped into a copper calorimeter of mass 100 g, containing 500 g of water. The temperature of the calorimeter is lowered from 30° to 11°.6 C. Determine the heat of fusion of water. Specific heat of ice = 0.5. *Ans.* 77.13 calories per gram.

4. What is the heat of fusion of water, expressed in British thermal units? *Ans.* 144 B. T. U. per lb.

5. How many inches of rain at 10° C must fall in order to melt a sheet of ice $\frac{1}{2}$ in thick? *Ans.* 3.664 in.

6. If 10 lb of water should freeze in a cellar containing 3000 cu ft of air, how much would the air be warmed, assuming the specific heat of air as 0.24 and the mass of one cubic foot of air as 0.08 lb? *Ans.* 13°.89 C.

7. What is the thermal capacity of one cubic foot of air in British thermal units per degree Fahrenheit? in calories per degree Centigrade?

Ans. (*a*) 0.0192 B. T. U. per degree F.
(*b*) 8.7 calories per degree C.

8. A loop of wire 0.02 cm in diameter is placed over a piece of ice, and a 4 kg weight is hung from it. The length of the wire in contact with the ice is 5 cm. Find the average pressure under the wire. At what temperature will the ice under the wire melt? *Ans.* −0°.29 C.

9. If 5 kg of water at 25° C be placed in a porous jar through which some water can gradually pass and evaporate, how much will have to evaporate in order to cool the remaining water 8° C below the temperature of the surroundings, assuming the water equivalent of the jar to be 50 g?

Ans. 68.2 grams.

10. How large a portion of water, undercooled to −12° C will freeze when crystallization takes place? Disregard the water equivalent of the vessel. *Ans.* 15 per cent.

11. How much heat is necessary to change 50 g of ice at −10° C to steam at 150° C? *Ans.* 37,205 calories.

12. An aluminium cup (specific heat 0.21) weighing 80 g contains 483.2 g of water at 18° C. Steam is passed into it until the temperature is raised to 37° C. The calorimeter with water weighs now 579.2 g. Calculate the heat of vaporization. *Ans.* 530.75 calories per gram.

13. How much would the air in a room 6 × 5 × 3 m be warmed by the condensation alone of 1 kg of steam in the radiator? *Ans.* 19°.26 C.

14. Find the correction which must be applied to a thermometer which gives the boiling point of water at 98°.5 C, when the barometer stands at 75 cm. *Ans.* +1°.13 C.

15. At what temperature will water boil in Denver, 5000 ft above sea level? *Ans.* 94°.88 C.

16. How much water vapor will be produced if 5 kg of water superheated to 105° C suddenly begin to boil under atmospheric pressure? The density of saturated water vapor at 100° C is 0.000606 g/cc. *Ans.* 76,681 cm^3.

17. How much water at 100° C and under atmospheric pressure can be vaporized by the burning of 1 kg of coal? Consider all the heat units produced as available for the evaporation of water. See problem 5, p. 232.
Ans. 14,498 grams.

18. A mass of saturated steam at 100° C is inclosed in a cylinder furnished with a frictionless piston of 400 sq cm area. No heat is supposed to leave the cylinder through the walls, and the vapor is allowed to do work by pushing the piston 10 cm out against a pressure of 75 cm of mercury. How much steam will be condensed? *Ans.* 0.177 grams.

19. The dew point of air at 25° C is found to be 17° C. Assuming the vapor tension of water at 17° C to be 14.4 mm of mercury, determine the relative humidity. What would be the relative humidity if the temperature of the air had been 20°? (See Arts. 201 and 214.)
Ans. (*a*) 61.2 per cent.
(*b*) 82.8 per cent.

20. How much heat is absorbed when 1 kg of liquid air is boiled under atmospheric pressure and subsequently heated to 20° C? Compare this amount with the heat absorbed by 1 kg of ice melted at 0° C and the water subsequently heated to 20° C. (Boiling point of air − 190° C; heat of vaporization of air 50 cal per gram.)
Ans. (*a*) 99,770 calories.
(*b*) 100,000 calories.

21. How much external work is done by 1 kg of water when it freezes at 0° C under atmospheric pressure. What would be the change in the heat of fusion of water if this amount of energy were not included?
Ans. (*a*) 9.304×10^7 ergs.
(*b*) 0.0022 calorie per gram.

22. How much external work is done by the transformation at 100° C of 1 kg of water into steam under normal pressure? How large would be the change in the heat of vaporization of water if this amount of energy were not included? (Density of steam at 100° C = 0.000606 g/cm^3; of water at same temperature, 0.959 g/cm^3.)
Ans. (*a*) $167,000 \times 10^7$ ergs.
(*b*) 39.9 calories per gram.

DISTRIBUTION OF HEAT

CHAPTER XXVIII

CONDUCTION

221. Three Modes of Distribution of Heat. In a general way we may speak of the transfer of heat from one body to another, and such transfer is always involved whenever bodies, originally at different temperatures, come to thermal equilibrium (Art. 148). This state is reached, in the case of two bodies thermally insulated from all other bodies, by a mutual approach toward some intermediate temperature. The temperature of one body rises and that of the other falls, by intervals which vary inversely as the thermal capacities of the two bodies in question.

In case one body be connected to some source of heat, the tendency is to maintain this body at some definite temperature and to bring surrounding bodies to the same temperature through the transfer of heat. In all cases, however, heat is transferred *from the body of higher temperature to the body of lower temperature.* From this point of view, difference in temperature is seen to be analogous to difference in level, or to difference in hydrostatic pressure in liquids, and to difference in pressure in connected reservoirs containing gases. In every case the difference in temperature or thermal pressure determines both the direction and the rate of transfer of heat.

Heat may be distributed in any one of three different ways, or, more generally stated, it may be distributed in all three ways at the same time. The three modes of distribution of heat are by *Conduction*, by *Convection* and by *Radiation.*

By *Conduction* is meant the flow of heat through an unequally heated body, or system of bodies, from points of higher to points of lower temperature. This mode of distribution of heat is exemplified in the heating of a metal rod by placing one end in a

Bunsen burner flame. The part in the flame soon becomes quite hot, the molecules of adjacent parts have their motion quickened through the impact of those in the hotter part, and a transfer of heat takes place to points of lower temperature. In this way is set up a steady flow of heat through a rod, between whose ends a definite difference in temperature or thermal difference of level is maintained. Such a rod is said to be a conductor of heat, and the relative ease with which such transfer is made is termed the thermal conductivity of the metal.

By *Convection* we mean the transference of heat by the bodily movement of heated particles of matter. In buildings heated by steam, by hot air or by hot water we find excellent examples of convection of heat.

By *Radiation* we mean the transfer of energy from point to point in space by means of waves set up in the ether. The earth is heated by radiation from the sun. The sunlight passing through the window pane brings both light and "heat" into the room. If we hold our hands above a heated stove, the hand is heated both by convection through the air and by radiation. If, however, we hold the hand at the side of the stove and at the same distance from it, the hand will still be heated, but in this case it will be due to *heating by radiation* only.

Distribution of heat by radiation is characterized by the absence of any temperature effect upon the medium between the hotter and the cooler body. It is true the hotter body loses heat, but from the definition of heat (Art. 146) it is apparent that the energy after it leaves the hotter body can no longer be called heat. It is "energy of radiation" and follows the laws of radiation. But when this form of energy is absorbed by a body, it is retransformed into heat, and the final result will be the gain of heat by one body at the expense of another body at a higher temperature.

From the foregoing it is clear that the so-called radiant "heat" is more closely connected with the subjects of light and electrical waves than with heat. For this reason radiation, including all three phenomena mentioned, will be treated in a later chapter.

222. The Temperature Gradient. It has been shown (Art. 147) how conduction of heat may be explained by the molecular theory of matter as an equalization of molecular kinetic energy. If a rod of metal be placed in a flame, the rise in temperature is at first largely influenced by the specific heat of the substance, since the smaller the specific heat, the more rapidly the temperature will rise.

After a short time, however, the temperature along the bar becomes constant, falling off more or less rapidly from the

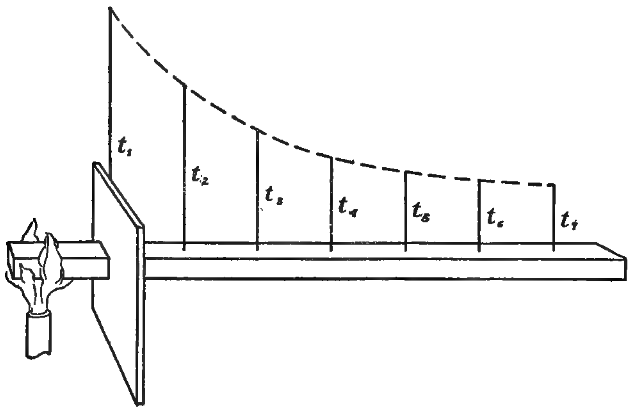


FIG. 112.

hotter to the cooler end (Fig. 112), while heat flows steadily through the bar at a definite rate.

The temperature gradient at a point *is the space rate of change of temperature*, or

$$gradient = \frac{t_1 - t_2}{l} \tag{294}$$

It is represented by the slope of the temperature curve.

Different bodies vary greatly in their ability to conduct heat. Silver is a very good conductor, iron not so good. A glass rod, placed in a flame, does not become too hot to touch more than a few centimeters from the flame, while we can hold a match until the flame almost touches the fingers.

The better thermal conductor a substance is, the smaller is the temperature gradient, other things being equal. or, the

smaller is the decrease of temperature for a given distance from the flame.

We may compare the relative conductivity of copper and iron by taking a copper and an iron wire of the same cross section, twisting their ends together, and attaching to them at short distances apart small pellets of wax. On placing the twisted ends in the flame of a Bunsen burner, the pellets will drop off as soon as the temperature has risen sufficiently to melt the wax. It will be seen that the distance through which the pellets have dropped off is much larger on the copper than on the iron rod. This shows that copper is a better conductor of heat than iron.

223. The Coefficient of Thermal Conductivity. Let two sides of a plate of thickness l be kept at constant temperatures t_1 and t_2, t_1 being larger than t_2. Heat flows through the plate from higher to lower temperature. The quantity of heat passing through any area A is proportional to this area, to the temperature gradient, and to the time τ during which the heat flows. Thus we obtain

$$H = kA\frac{t_1 - t_2}{l}\tau \tag{295}$$

where k is a proportionality factor depending upon the material of the plate. It is called *the coefficient of thermal conductivity* and may be defined as *the time rate of heat conduction per unit area per unit temperature gradient.* It is numerically equal to the heat transferred in unit time through unit area of a plate of unit thickness, if unit difference of temperature be maintained between its two faces.

The numerical value of k depends upon the units chosen for the other quantities. It is said to be expressed in C. G. S. units if H be given in calories, $t_1 - t_2$ in degrees Centigrade and the rest of the quantities in C. G. S. units. The experimental determination of k is very difficult, chiefly on account of the loss of heat from the sides of the body through which the heat passes.

224. Conduction of Heat in Liquids and Gases. In general liquids are poor conductors of heat. Water in a test tube may

be boiled in the upper part of the tube, while the lower part contains a piece of ice held down by a piece of wire gauze.

Gases are even poorer conductors than liquids. Porous substances such as wood, wool and asbestos are poor conductors, on account of the large amount of air enclosed in the interstices between the solid material. Such substances are much used to prevent loss of heat from steam pipes, fireless cookers, etc. For this same reason loose clothing is warmer than snugly fitting garments.

In all determinations of the conductivity of fluids it is necessary to avoid currents in the fluid, since they would produce an equalization of temperature by convection instead of by conduction.

225. The Leidenfrost Phenomenon. If a drop of water be carefully placed by means of a pipette upon a metallic plate, which has been heated to low red heat, the water does not boil, but forms a flattened spheroid, rolls about the surface, evaporating quietly, because a thin film of vapor is formed between the plate and the drop. Owing to the low conductivity of the vapor, heat enters the water but slowly and the drop is at a temperature several degrees below the boiling point. This is called the *Leidenfrost phenomenon*, or the phenomenon of the *spheroidal state*.

If the plate and the drop be connected by an electric circuit containing a bell, the bell does not ring, since there is no contact between plate and drop. When the plate is allowed to cool again, the water makes contact with the plate, boiling sets in, and the bell begins to ring.

TABLE XVI

COEFFICIENTS OF THERMAL CONDUCTIVITY IN C. G. S. UNITS

SUBSTANCE	k	SUBSTANCE	k
Silver	1.00	Glass	0.0015
Copper	0.90	Ice	0.005
Zinc	0.25	Wood	0.002
Iron	0.15	Water	0.0015
Mercury	0.016	Air	0.00005

226. Applications. Mention has already been made (Art. 148) of the fact that metals appear colder in cold weather and warmer in warm weather than wood, wool or similar substances. This is easily explained by the difference in thermal conductivity of the substances in question.

A flame does not actually touch a body which is kept at a temperature much lower than that of the flame. Thus water may be boiled in a paper tray over the flame of a Bunsen burner, since the paper remains approximately at the temperature of the boiling water. If we place a piece of wire gauze a few centimeters above a Bunsen burner, turn on and ignite the gas above the gauze, we shall see that the flame will not "strike back" so long as the temperature of the gauze is kept low by conduction of heat away from the flame. If the gauze be lowered over a flame, the flame will not strike through for some time and will do so only after the gauze has become quite hot.

This fact is utilized in Davy's safety lamp. A fine wire gauze entirely surrounds the flame of the lamp. If there be any explosive gases in the mine, they will pass through the gauze into the lamp and ignite there, burning with a bluish flame. For some time the gauze remains cold enough to prevent an explosion of the gases outside. The burning of the gases in the lamp is a warning to the miner to leave the workings until ventilation has restored the atmosphere to a safe condition.

CHAPTER XXIX

CONVECTION

227. Cause of Convection. If a vessel filled with water, in which some solid particles float, be heated, the solid particles show by their motion that a current rises from a point directly over the flame and flows downward again along the cooler walls of the vessel. The upward motion is due to the expansion and decrease of density of the water as the temperature rises. This raises the surface over the heated portions, and the water flows to the lower level above the cooler portions, thus increasing the pressure. The cooler water flows then from points of higher pressure to those of lower pressure and drives the heated portions upward.

In the same manner the density of air is decreased by heating; the light air expands, flows off at the sides, and the cold air presses into the areas of low pressure, causing an upward motion of the heated air. Convection currents in air may often be observed in the "shimmering" of the air over heated plains. The same effect can be shown by passing a beam of light through the air above a heated plate. The heated air is lighter, and has a different index of refraction from that of the cooler air around it, and its movement can be distinctly seen on a screen.

Convection currents do not, strictly, transfer heat, but heated particles. An actual transfer of heat occurs when these particles come into contact with cooler bodies and heat them by conduction.

228. Convection in Liquids. Convection of heat by liquids is employed in the heating of buildings by hot water (Fig. 113). The water rises from the boiler B through a system of pipes to the rooms, loses heat in the radiators R_1, R_2, and is led back to the lower part of the boiler by the return pipes. Every hot-

water system must be supplied with an open tank, *T*, to allow for the expansion of the water when heated.

Some ocean currents, originating near the equator, have been considered as convection currents. The level cf the ocean rises with increase of temperature, and the water flows towards the lower levels farther north, being replaced by cold water from below. Wind, however, is a more important factor in directing the ocean currents, than the extremely small expansions produced by the heating of the ocean.

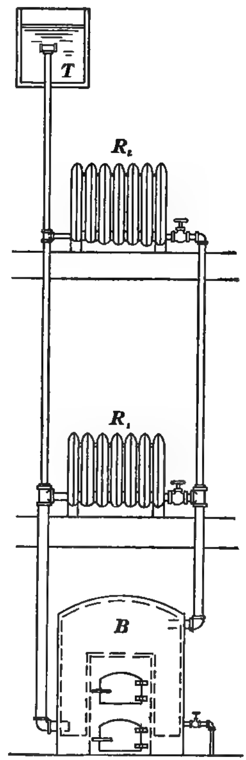


Fig. 113.

229. Convection in Gases. Convection currents of air are in common use for heating buildings and for ventilation. Convection currents are very efficient equalizers of temperatures. The amount of heat carried in this way is quite large. Since one cubic meter of air at 20° C weighs 1.2 kg and its specific heat is 0.24, a cooling of this amount of air to 15° C would release 1440 calories.

The following experiment shows clearly the effect of convection currents of different gases upon a heated body. Inclose a fine straight platinum wire in a glass tube so that it hangs vertically in the axis of the tube. Arrange the experiment so that the tube can be exhausted. Heat the wire to dull red heat by a definite electric current. Upon exhausting the air the wire will lose much less heat than before, owing to the absence of convection currents, and its temperature will be raised to a bright yellow heat. Now let the tube be filled with hydrogen gas. The same electric current as before will not be sufficient to produce a glow in the wire, showing that the convection currents of the hydrogen have a much larger cooling effect than those of air.

The bulbs of incandescent lamps are exhausted and thus loss of heat by convection currents from the glowing filaments to the glass envelope is avoided. The so-called "Dewar flasks" (Fig. 114) are double-walled glass vessels, the space between the walls being an extremely low vacuum, and only a very small heat exchange takes place between the inside and outside of the vessel. Liquids with low boiling points, such as liquid air or liquid hydrogen, may be kept much longer in Dewar flasks than in ordinary vessels, since the rate of evaporation is greatly reduced, on account of the slowness with which the heat needed for evaporation passes to the liquid. Recently such flasks have been placed on the market under the name of "Thermos-bottles," and serve equally well either for keeping a liquid hot or for keeping it cold.

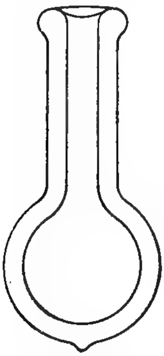


FIG. 114.

230. Convection Currents in the Atmosphere. Land and sea breezes are examples of convection currents. The heating of the land during the day produces an expansion of the air, an overflow at higher altitudes, and consequently a decrease in pressure. The air over the sea, being now at a higher pressure, flows towards the region of low pressure and forces the heated, lighter air upwards. This inflow from the sea is called *the sea breeze.* At sunset the opposite occurs, since the land cools rapidly, while the ocean, owing partly to the large specific heat of water, remains at nearly the same temperature as during the day.

The trade winds furnish examples of convection currents on a larger scale. The heated air in the equatorial region rises while cooler air flows in on the surface from the north and the south. On account of the rotation of the earth these cool *trade winds* come from an easterly direction. The air which has risen in the tropics flows off towards the poles and descends again to the surface of the earth at a latitude of about 35°. Now since these warm winds descend from greater heights and have therefore a larger velocity toward the east than the surface of the earth, they

will flow from the southwest in the northern hemisphere and from the northwest in the southern hemisphere.

The atmospheric disturbances, due to these causes, seldom reach an elevation of more than 3000 m or two miles. In this region cloud formation and precipitation takes place, while the temperature variations are usually quite irregular. Above this lowest region of terrestrial disturbance there is another region, extending to about 11,000 m, in which the temperature in general decreases uniformly with height to about $-55°$ C in middle latitudes, and in which the motion of the atmosphere is in an easterly direction. This region is comparatively free from condensation. The cirrus clouds are found in its uppermost portion.

Above the elevation of 11,000 m there is another distinct region, called *the isothermal region*, from the fact that here the temperature changes but slightly with the elevation, however much it may differ from place to place and from day to day, the average for middle latitudes being $-55°$C. Vertical convection currents of the air, producing adiabatic expansion and cooling, are impossible in this region. It has been explored to a height of 29,000 m by means of balloons, carrying registering instruments. At the height of 70,000 m there appears a rapid change with elevation in the composition and density of the atmosphere. This conclusion is drawn from a study of twilight and other phenomena, all of which support the theoretical deduction that our atmosphere at an elevation above 70,000 m consists mainly of hydrogen and helium, and at lower levels chiefly of nitrogen.

Problems

1. Water is boiled in an iron vessel having a heating surface of 400 cm^2 and a thickness of 4 mm. How much water will be evaporated per minute, if the surface exposed to the fire be kept at 280° C? *Ans.* 3.01 kilos.

2. A lake having a surface area of 9000 square meters is covered by a sheet of ice 5 cm thick. How much heat will pass through the ice in two hours, if the temperature of the air be $-10°$, and the ice do not appreciably increase in thickness? *Ans.* 648×10^7 calories

3. The walls of a certain refrigerator have an area of 10,000 cm^2, are 8 cm thick, and are made of wood. Find how much ice may be expected to melt in a day, if the outside temperature be 25° C? *Ans.* 67.5 kilos.

4. How much coal must be burned to compensate for the loss of heat due to conduction for one day through a glass window 4 mm thick and having an area of 2 square meters, supposing the air in the room next to the glass to be at 25° C, and the outside air at − 10° C? Why is this amount much greater than that actually needed? *Ans.* 29 kilos.

5. How much heat would be lost per square decimeter per minute by a man clothed in a fabric 0.3 cm thick, having a coefficient of conductivity equal to 0.00012 C. G. S. units, assuming the temperature of the air to be 5° C, and the temperature of the body 30° C? *Ans.* 60 calories.

6. A balloon of nearly spherical shape and of a capacity of 1000 cubic meters is filled with air of a temperature of 30° C above that of the outside air, which is at 20° C. What is the force driving the balloon upward? *Ans.* 10.965×10^7 dynes.

ELECTRICITY AND MAGNETISM

MAGNETISM

CHAPTER XXX

ACTION-AT-A-DISTANCE THEORY

231. Magnets. A certain iron ore, called magnetite or loadstone, has the characteristic property of attracting iron filings. The same property may easily be given to a rod of steel by rubbing it repeatedly with the loadstone from one end to the other, always passing in the same direction along the rod. The steel thus treated is said to have been *magnetized*, and the rod is called *a magnet*. Substances which are attracted by a magnet are called magnetic substances.

A magnet when suspended by a thin untwisted thread also shows the characteristic property of assuming a definite orientation with respect to the geographical meridian. Thus a long, thin magnet, or magnetic needle, if undisturbed by mechanical forces and uninfluenced by other magnets, or magnetic substances, always places itself in an approximately north-south direction. The end of the needle pointing towards the north is called the *north-seeking pole* or the *positive pole*, the one pointing towards the south, the *south-seeking pole* or the *negative pole*. Frequently the shorter expressions *north pole* and *south pole* are used. For a more accurate definition of a pole, see Art. 234.

232. Mechanical Forces between Magnets. A magnetic needle suspended by a thread or mounted upon a sharp point is deflected when another magnet is brought near it, the direction of the deflection showing in every case that *like poles repel and unlike poles attract each other.* Thus the north pole of the

needle is repelled by an approaching north pole and attracted by a south pole.

The quantitative expression for the mechanical force produced by the mutual action of two magnetic poles was first given by Coulomb[1] in 1785: *The force of attraction or repulsion between two poles is inversely proportional to the square of the distance between the poles and directly proportional to the product of their pole strengths.* Denoting the proportionality factor by k, Coulomb's law may be written:

$$F = \pm k \frac{m_1 m_2}{d^2} \tag{296}$$

The quantities m_1 and m_2 are called the *pole strengths* and are characteristic properties of the two poles. The force is considered as positive in the case of repulsion, and negative in the case of attraction. The mechanical forces due to magnetic action are enormously larger than the force of attraction due to gravitation between the masses of the magnets.

233. The Action-at-a-distance Theory. Coulomb's law is identical in form with the law of gravitation, and it was only natural that the first theory of magnetism should be an exact duplicate of the theory of gravitation, as held at that time. It was assumed that magnetism was a substance, and that quantities of magnetism, represented by m_1 and m_2 in Coulomb's law, had the innate power of attracting or repelling other quantities of magnetism separated from them in space.

In order to explain attraction as well as repulsion, it was necessary to assume the existence of two different kinds of magnetism of opposite nature, *positive* and *negative*. Magnetism was also assumed to be an imponderable substance, since the magnetization of a piece of iron or steel did not change its weight. We shall see later that this theory is unsatisfactory, and that the phenomena in question may be explained much better by the assumption that the medium between the poles is the real seat of magnetic action.

However, Coulomb's law is independent of any interpretation

[1] Coulomb, *Mem. de l'Acad.*, 1785, p. 603.

which may be given to the quantities m. It has been shown to be exact by numberless experiments, and can be used directly for the solution of problems. We shall, therefore, for the time being, use the term *pole strength* as if it denoted a definite quantity of magnetism. This quantity of magnetism may be defined as an hypothetical substance, which, when placed upon a body, renders it a magnet, and by its action at a distance causes the attraction or repulsion manifested between magnets. We shall also derive the concepts of some other magnetic quantities, as they have been developed by the action-at-a-distance theory. The newer theory will be given later.

234. Poles of a Magnet. If a magnet be dipped into iron filings and withdrawn, the filings are seen to cling to it, being crowded together in dense masses at the ends, but decreasing in amount from the ends toward the middle, where none adhere. This seems to show that the magnetism is distributed over the surface of the magnet, being most dense at the ends. In very thin magnets the magnetism is concentrated at points very near the ends, and almost no iron filings are seen to adhere along the sides.

At some distance from a magnet the mechanical forces acting upon a small, thin magnetic needle may be considered as proceeding from two points in the magnet, where we may assume all the magnetism to be concentrated, just as the effect of gravitation may be considered as proceeding from the center of gravity of a large mass rather than from each individual particle of matter.

When a small compass needle is brought into the neighborhood of a large magnet, it takes up a definite position, which is determined by the resultant of the forces acting upon the poles of the needle. Thus (Fig. 115), at the point P the force on the north pole of the compass needle is directed towards Q. This force PQ is the resultant of the forces Pn and Ps acting according to Coulomb's law between the north pole of the needle and two definite points in the larger magnet, N and S. If the compass needle be not brought too close to the magnet, the points N and S have always the same position in the mag-

net for any position of the needle, and are called the *poles of the magnet.* The straight line *NS* connecting the poles is called the *axis of the magnet.*

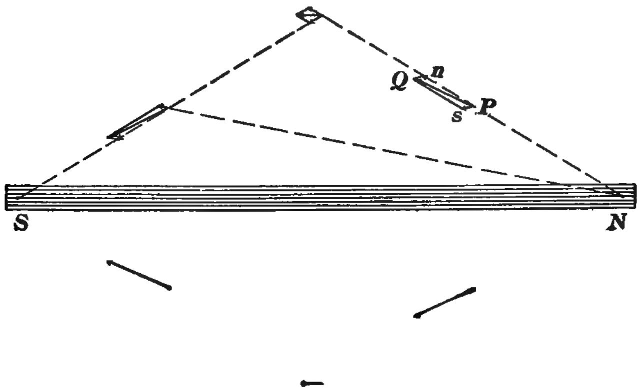


FIG. 115.

When the magnitudes and directions of the forces are carefully determined, it is found that *the two poles of a magnet are always of the same strength* and are, in general, at equal distances from the ends of the magnet. In a long, thin magnet the poles are quite near the ends, but the distance between the poles is always less than the length of the magnet.

235. Unit Pole. Coulomb's law enables us to select a unit of pole strength. All magnetic units are based upon the C. G. S. system. If we make the mechanical force F in equation (296) one dyne, the distance between the poles one centimeter, and agree that k shall be unity when the poles are placed in a vacuum, we obtain *unit pole strength* or *unit pole.* Hence, unit pole is *that pole which at unit distance in vacuo from an equal and similar pole repels it with a force of one dyne.* The force of attraction between two unit poles of opposite sign is evidently also one dyne. No specific name has been given to the unit of pole strength.

In magnetic theory it is frequently of advantage to consider the effect of a single pole, and while it is impossible to obtain a single pole, yet the poles of very long and thin magnetic

needles may be considered as approximately separate poles, since, owing to the length of the magnet, the effect of the remote pole is practically negligible.

236. Intensity of a Magnetic Field. The space surrounding a magnet is called a *magnetic field*. Coulomb's law gives an expression for the mechanical force produced by the mutual action of two magnetic poles, expressed in terms of their pole strengths. We may, however, express the force acting upon a single pole without reference to the strength of any other pole. This is exactly analogous to the two ways of expressing a force in mechanics. Although we know that a gravitational force can exist only between *two masses*, yet the force acting upon a given mass M' at a given point, is expressed by the equation

$$F = aM'$$

where a is the acceleration at that point due to any gravitational field whatever, without any reference to a second mass.

Similarly, if a force act upon a magnetic pole, it is not necessary to know the exact position and strength of any other pole; we may express it as

$$F = Hm \tag{297}$$

where m is the pole strength of the magnet and F the mechanical force acting on each pole. The proportionality factor H is called the *intensity of the magnetic field*, just as the acceleration a may be called the intensity of the gravitational field.

The *intensity of the magnetic field at a point is*, therefore, *the force per unit pole acting at that point.* It is numerically equal to the force acting upon unit pole. It is a vector quantity, lying in the same direction as the force acting upon a positive pole. The unit of magnetic field intensity is *one dyne per unit pole*, and is called the *gauss*, after the famous German physicist, Gauss (1777–1855).

Of course there exists no force at the point in question and, from the point of view of the action-at-a-distance theory, the intensity of the field has no physical meaning, unless a magnetic pole be placed at the point. We can nevertheless say that the

intensity of a magnetic field at a given point has a definite value, just as we say that the acceleration due to gravity has a definite value at a given point, and that it is independent of the presence or absence of a mass at that point.

A measurement of the intensity of the magnetic field at a point requires that a pole be brought to the point, and that the mechanical force exerted upon this pole be determined. A magnetic field is said to be *uniform* when its intensity at every point is the same in magnitude, direction and sense. The direction of the field is the same as the direction of its intensity.

237. Magnetic Moment. If a magnet be placed in any uniform field of intensity H (Fig. 116), the force on each pole is Hm, and the two parallel, equal and opposite forces acting upon the magnet constitute a couple (Art. 42) tending to rotate the magnet. Let the axis of the magnet make an angle α with the direction of the field intensity, and let l be the distance between the poles. The moment of the couple, or the torque $\mathcal{T}$, acting upon the magnet, is then

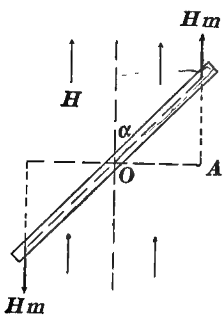


FIG. 116.

$$\mathcal{T} = 2\,Hm\,\frac{l}{2}\,\sin\alpha \qquad (298)$$
$$= Hml\sin\alpha = HM\sin\alpha$$

where M is called the *magnetic moment* of the magnet. The magnetic moment of a magnet is therefore *the product of the strength of one of its poles into their distance apart.* It is numerically equal to the moment of the couple acting upon the magnet when placed with its axis at right angles to a uniform field of unit intensity. The unit of magnetic moment is one dyne centimeter per gauss.

238. Permeability. Coulomb's law (Art. 232) contains a constant k whose value in vacuo was chosen as unity. Accurate measurements show that the force of attraction between two poles varies with the medium between the poles, being in some cases larger, in others smaller, than it would be in vacuo.

In order to take this experimental result into account, the constant k cannot, in general, be assumed to be unity, but must depend upon the medium. For reasons which will appear later the expression $\frac{1}{\mu}$ is substituted for k, where μ, a characteristic property of the medium, is called the *permeability* of the medium.

Coulomb's law is now written

$$F = \pm \frac{1}{\mu} \frac{m_1 m_2}{d^2} \tag{299}$$

By combining equations (297) and (299), the field intensity H, at a point distant d cm from a single pole m_1, is found from the force exerted there upon another pole m_2, to be equal to

$$H = \frac{F}{m_2} = \pm \frac{1}{\mu} \frac{m_1}{d^2} \tag{300}$$

Consequently, the permeability of a given medium is *a property modifying the action of magnetic poles immersed in this medium.* The permeability of gases and liquids differs but little from unity. For air at 20° under atmospheric pressure it is 1.000005, and for alcohol 0.999990.

239. Magnetism a Molecular Property. If a magnet, as a magnetized watch spring, be broken into small pieces, each piece is found to be a complete magnet, possessing both north and south poles. However far the subdivision may be carried, it is impossible to obtain a piece which contains but a single pole. The conclusion seems reasonable that each molecule of a magnet or of any magnetic substance is itself a complete magnet, and that the magnetic effect produced by the original magnet was simply the resultant effect of these molecular magnets. Though this theory does not explain the nature of magnetism, it helps us to understand certain well-known phenomena. Thus, in the case of an iron rod which has been magnetized by rubbing it with a magnet, it may be assumed that the magnetization so produced is due to a turning of these individual molecular magnets under the influence of the magnetizing pole. Before treatment, these molecular magnets were without definite orientation, and their external magnetic effect was zero. Under the influ-

ence of the magnetizing pole they have been forced into magnetic alignment, with like poles all turned in the same direction. Their external magnetic effect may now be considerable; they form a magnet.

In hardened steel these molecular magnets are brought into alignment with greater difficulty than in soft iron, owing to greater internal friction. But the steel rod, once magnetized, retains its magnetism for the same reason, and has become a permanent magnet. On the other hand if a magnetized rod be hammered or twisted or subjected to a mechanical shock, it loses part of its magnetism, since some of the molecules return to their natural positions.

If a glass tube filled with iron filings be placed in a strong magnetic field, produced by an electric current (Art. 319), the tube of filings will be found to be a magnet. It loses its magnetic properties, however, when the filings are thoroughly shaken. The experiment may be varied by using a short, thick tube, having its ends closed by glass plates, and containing fine iron filings suspended in glycerine. Place the tube lengthwise in a parallel beam of light. No light will pass through the tube. But if the iron filings be magnetized by a magnetic field parallel to the axis of the tube, the particles arrange themselves with their axes in this direction, and permit a considerable part of the light to pass. As soon as the magnetizing influence ceases, the light is again cut off.

240. Loss of Magnetic Quality at High Temperatures. We have seen that iron is attracted by a magnet. The iron is said to possess magnetic quality. The molecular theory of magnetism explains this on the assumption that the particles of iron are turned under the influence of the magnetizing field. The iron itself becomes for the time being a magnet, or *magnetism is induced in the iron.* At red heat, iron loses its magnetic quality. This not only means that a magnet may be completely demagnetized by being heated to about 800° C, but *further*, that iron, when heated to this temperature, *is no longer capable of being attracted by a magnet.* On cooling, however, the iron regains its magnetic quality at a somewhat lower temperature. From the

point of view of the molecular theory of heat (Art. 147), the motion of the particles at high temperatures is so much increased that they cannot be kept "lined up" by the influence of the magnetic field.

Nickel and cobalt also possess magnetic quality, but lose it at much lower temperatures than iron.

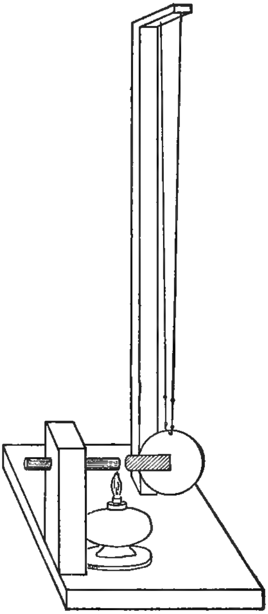


FIG. 117.

The following experiment shows the loss of magnetic quality at high temperatures. A flat strip of nickel is soldered to a blackened copper disk and suspended by silk threads, forming a pendulum (Fig. 117). A permanent magnet is placed a short distance from the nickel so that it holds the disk deflected through a small angle. When the nickel is heated by an alcohol lamp, its temperature rises, and it soon loses its magnetic quality. Being no longer attracted by the magnet, the pendulum swings away and speedily cools, owing to the rapid radiation from the blackened copper disk. Upon its return it is again attracted to the magnet, since the temperature has fallen below the point at which nickel regains its magnetic quality.

CHAPTER XXXI

THE ETHER-STRAIN THEORY

241. Deformations in Elastic Bodies. Faraday explained magnetic phenomena by the assumption of elastic deformations of the ether. Although it is impossible to imagine an ordinary body whose elastic properties correspond in every detail to those of the ether, yet it will be helpful to derive the fundamental concepts of the ether-strain theory through a comparison with the phenomena observed in ordinary elastic bodies.

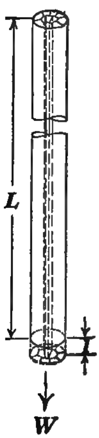


FIG. 118.

Thus, consider a homogeneous cylindrical wire of length L, fastened at one end and stretched by the weight W of a mass M hung from its lower end (Fig. 118). The wire is strained, being lengthened by l, and at the same time a stress appears in the wire. The strain is measured by the relative lengthening of the wire $\frac{l}{L}$, while the stress is measured by the force per unit area tending to restore the wire to its original unstrained condition. Since the lengthening at every point in the wire is proportional to the length measured from the upper end, the strain is the same throughout the wire.

To fix our ideas, let us assume that the strain is P. This strained condition of the wire may be represented by assuming the cross section of the wire to be divided into small equal areas such that P of them cover one square centimeter. By drawing through every point of the boundary lines of these small areas lines parallel to the surface of the wire, the wire is divided up into a number of tubes which pass through its whole length. The strain at any point may then be measured by the *number of tubes* passing through unit cross section.

If the material, instead of being cylindrical, have the shape of a cone, and if a force be applied at one end (Fig. 119), it is evident that the cross section of the strain tubes must increase, and their number per unit cross section must decrease, as we pass from the apex to the base of the cone.

The internal force upon any area A of the cross section is the *product of the stress into the area*, or

$$F = A \times stress \tag{301}$$

and the stress is proportional to the strain, or

$$stress = M \times strain \tag{302}$$

where M is the coefficient of elasticity, in this case Young's modulus, and is a characteristic property of the substance. The larger the coefficient of elasticity, the larger will be the force needed to produce a given strain.

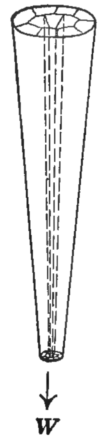


FIG. 119.

242. Magnetic Induction and Intensity of Field. In the Faraday-Maxwell theory it is assumed that the ether in a magnetic field is subjected to an elastic strain, which gives rise to a stress, tending to reduce the strain. This strained condition of the medium shows itself in the mechanical force existing between magnetized bodies. Such bodies always tend to move so as to reduce the strain in the ether. The stress at any point in a magnetic field may therefore be measured by the mechanical force acting upon a magnetic pole placed at that point, and it is generally agreed to call the magnetic stress unity if the force exerted upon unit pole be one dyne. Since by equation (297)

$$F = Hm$$

it is clear that the stress at any point in a medium is measured by the intensity of the field at that point.

As in the case of Hooke's law for elastic bodies (Art. 57), we have seen that the strain is proportional to the stress, so now, in magnetic relations, the magnetic strain B is proportional to the magnetic stress H, and is found by multiplying the field in-

tensity H by the constant μ, which is a characteristic property of the medium, and is called the permeability of the medium. The strain B, given by the equation

$$B = \mu H \tag{303}$$

is called the *magnetic induction.*

We see, therefore, that at any point in a magnetic field the field intensity corresponds to the *stress in an elastic body*, the *induction to the strain*, and *the reciprocal of the permeability to the coefficient of elasticity.*

243. Tubes and Lines of Induction. Lines of Force. Using Faraday's method of representing the strained condition of a medium by strain tubes, the whole field about a magnet may be considered as filled by tubes of induction, which, by their number per unit area, measure the strain of the medium.

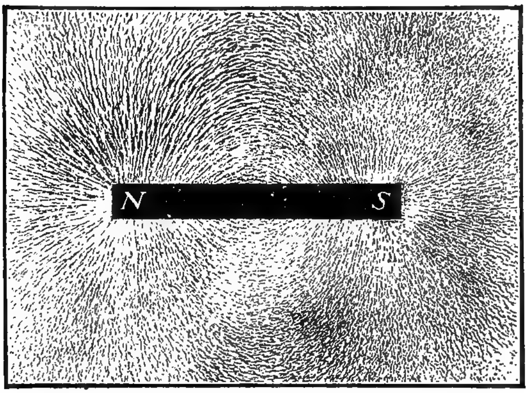


FIG. 120.

In a qualitative way this condition may be shown by sprinkling iron filings upon a sheet of paper spread over a magnet. On tapping the paper gently, these filings arrange themselves in curves, passing from one pole to the other (Fig. 120). The tangents to these curves give the direction of the forces acting upon the iron filings, and they show the general form of the strain tubes in the ether about the magnet.

In order to simplify the representation of the field, we may draw a single line for each tube, assuming each line to represent the axis of a tube. We thus represent the magnetic strain by *lines of induction* or lines of force, as they are frequently called. But since no force exists unless a magnetic pole be brought into

the field, and further, since the induction is entirely independent of the presence or absence of a pole at the point in question, we shall use in this text only the expression, "*lines of induction.*"

In order to distinguish between north-seeking and south seeking poles in this method of representing magnetic action, it has been agreed that *the lines of induction come out of a magnet at the north-seeking pole and enter at the south-seeking pole.*

No specific name has been given to the unit of induction. Since induction is measured by the number of lines per unit area, its unit may be taken as *one line per square centimeter* in a plane perpendicular to the lines. The total number of lines of induction through a given area is called *the magnetic flux.* For example, suppose the force exerted upon a unit pole placed at a point in a uniform field of permeability 1.5 to be 10 dynes. The magnetic intensity of the field at that point is 10 gausses, the induction is 15 lines per square centimeter and the direction of the lines of induction is the direction of the force acting upon the pole. The magnetic flux through an area of 20 square centimeters is 300 lines.

244. Properties of Lines of Induction. If we place two poles of opposite sign near each other, and render visible the lines of induction by means of iron filings (Fig. 121), it will be seen that the lines pass directly from the north pole of one magnet to the south pole of the other. The attraction between unlike poles may therefore be considered as the result of a tendency of the lines to shorten.

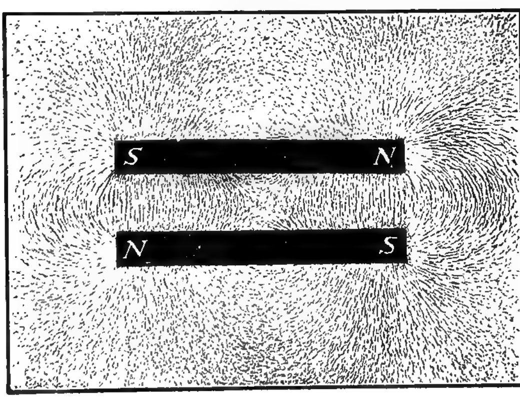


Fig. 121.

If we repeat the experiment with poles of like sign (Fig. 122), we shall see that none of the lines from either pole enter the other of like sign. The repulsion between like poles is therefore produced by a tendency of the lines to repel each other. These experiments indicate that in a magnetic field there exists a *tension in the direction* of the lines of induction and a *pressure at right angles* to this direction.

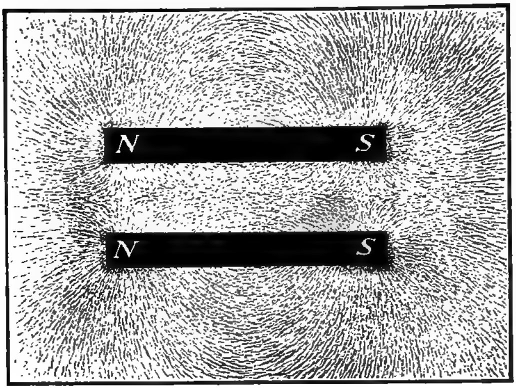


FIG. 122.

Since the intensity of the field H, at a distance of one centimeter from a single pole, is (Art. 238)

$$H = \frac{m}{\mu} \text{ gauss} \tag{304}$$

and

$$B = m$$

and since the field is symmetrical in all directions, there are by definition m lines of induction through every square centimeter *at one centimeter distance* from the pole, and $4\pi m$ lines through a sphere of unit radius or through any closed surface surrounding the pole. This means that 4π lines of induction proceed from unit pole. The number of lines of induction leaving or entering a pole is therefore independent of the surrounding medium.

245. Lines of Induction through a Magnet. It has been shown that the pieces of a magnet are always complete magnets. However far the subdivision of a magnet may be carried, lines of induction will always pass *into* the small pieces and *out of* them. From this the important conclusion must be drawn that "*the lines of force*" (or lines of induction) "*are closed curves*, passing in one part of their course *through the magnet*

and in the other part *through the space about it.* These lines are identical in their nature, qualities and number, both within the magnet and without."[1]

This assumption that the lines are continuous also explains the experimental fact, already mentioned (Art. 234), that the two poles of a magnet are always of the same strength.

246. Induced Magnetism. The lines of induction between the two arms of a horseshoe magnet run very nearly parallel to each other (Fig. 123), and the field between the arms is approximately uniform. This arrangement of the lines is completely altered, however, if a small piece of iron be placed in this field (Fig. 124). It is seen that the lines of induction crowd together into the iron. This shows that iron offers less resistance to the magnetic flux than does air, or that it is more *permeable.* It is this effect which has given the quantity μ its name. Substances whose permeability is greater than unity are said to be *paramagnetic*, or simply magnetic.

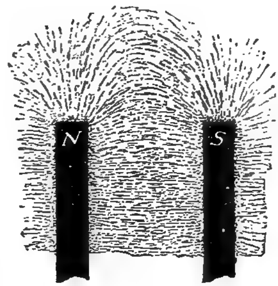


FIG. 123.

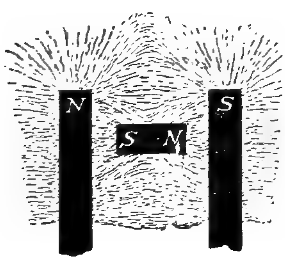


FIG. 124.

From the above figures it is seen that the lines coming from the north-seeking pole crowd into the iron at its nearer end, and the same number issue from the other end. Since that end where the lines go into the iron is a south-seeking pole, and that from which they come out is a north-seeking pole, we may also say that, upon the approach of a magnetic pole to a piece of iron, magnetism of the opposite sign is induced at the end of the iron nearest the pole, and an equal amount of magnetism of the same sign appears at the farther end of the iron. The same must be true

[1] Faraday, *Researches*, vol. iii, p. 417.

for any other substance whose permeability is greater than that of the surrounding medium.

But if the permeability of the body introduced into a magnetic field be less than that of the surrounding medium, it offers a greater resistance to the magnetic flux than does the original medium. In this case repulsion between the magnet and the body will result, and a needle made of such a substance will place itself at right angles to the magnetic field. In this position the resistance to the magnetic flux is a minimum. Substances whose permeability is less than that of air are said to be *diamagnetic*.

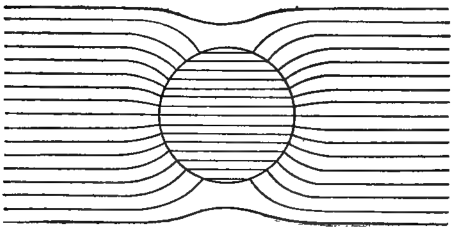


FIG. 125.

It may also be shown that feebly magnetic bodies when suspended in a more permeable medium behave as if they were diamagnetic. A small glass tube filled with a weak solution of ferric chloride, when placed in a strong magnetic field *in air*, will set itself parallel to the lines of induction, thus showing that it is *paramagnetic*. But if it be suspended in a more concentrated solution of ferric chloride, it will place itself at right angles to the magnetic field, as if it were *diamagnetic*.

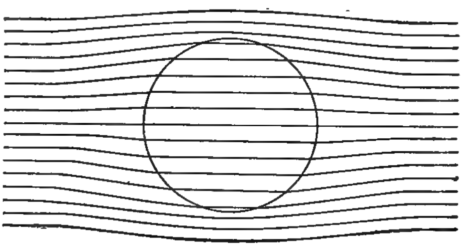


FIG. 126.

Figs. 125 and 126 show the distribution of the lines of induction in a sphere whose permeability is larger than that of the surrounding uniform medium in the first case, and smaller than that of the medium in the second case.

Substances whose permeability is very large, such as iron, nickel, and cobalt, are often called *ferromagnetic*. To this

group belongs an interesting alloy, called after its inventor, "Heusler's alloy," which contains no ferromagnetic substances, but is nevertheless strongly magnetic. It is an alloy of copper, manganese and aluminium. Recently several other such alloys, all containing either manganese or chromium, have been found to possess magnetic properties.

CHAPTER XXXII

MAGNETIC FIELD OF THE EARTH

247. The Earth a Magnet. A magnetic needle, if suspended so as to move freely, assumes a definite orientation. If the needle be placed on a cork on water, a rotation will be observed. This simple rotation can only be due to the action of a pair of equal, parallel and oppositely directed forces acting upon the two poles of the needle. Each force is defined by the equation

$$F = Hm$$

and since the two poles of the needle are equal, it follows that the intensity of the magnetic field of the earth is practically uniform at any given place (Art. 237).

The position which a magnetic needle assumes under the influence of the earth's field shows that the lines of induction on the surface of the earth follow in general a south-north direction, and that there is a magnetic *north pole* in the southern hemisphere, a magnetic *south pole* in the northern hemisphere.

248. Magnetic Declination. A magnetic needle does not, in general, place itself in the geographic meridian, but makes a small angle with it. This angle is called the *magnetic declination*, and varies from place to place. This fact was discovered by Columbus in 1492. The declination for Ann Arbor is at the present time very nearly 2° to the west.

Lines drawn on maps so as to connect all points of equal declination are called *isogonic lines* (Fig. 128). At present the line of zero declination passes in the United States from Charleston, through Asheville, Cincinnati, Fort Wayne and Lansing. For all points east of this line the declination is towards the west; for all points west of it the declination is towards the east. If drawn on a large scale, the isogonic lines

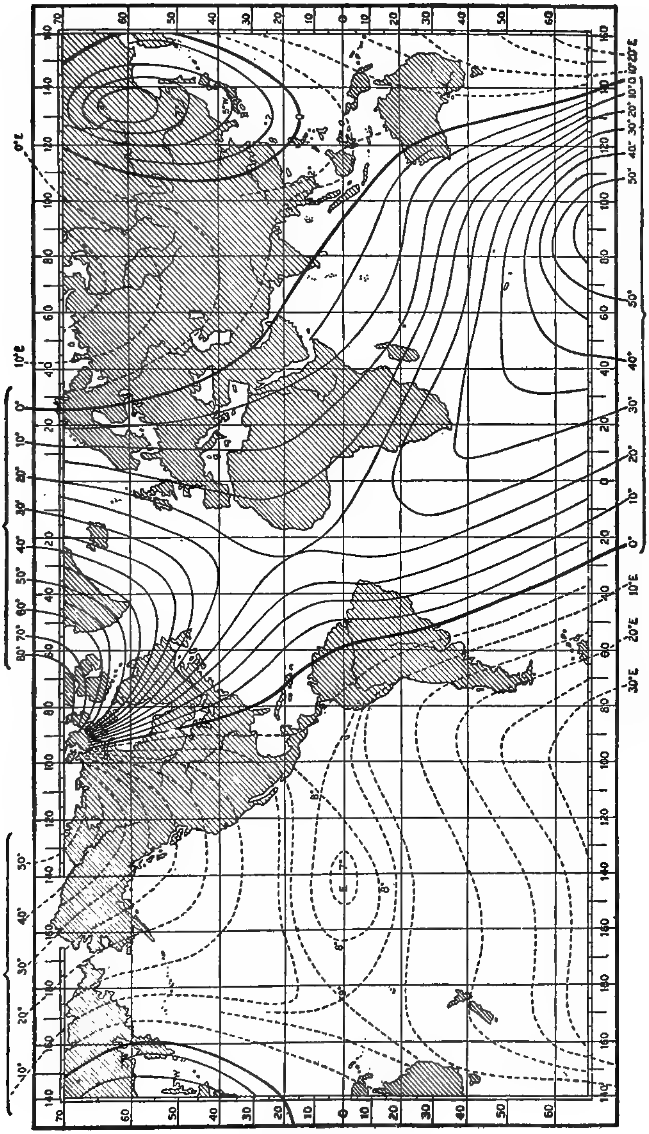


FIG. 128.

are by no means smooth curves, but show considerable irregularities.

249. Magnetic Dip. In 1576 Norman discovered that a magnetic needle, when supported at its center of gravity and free to turn around a horizontal axis, is not horizontal, but is inclined towards the horizon. In the northern hemisphere the north pole, and in the southern hemisphere the south pole, is depressed. The angle which such a "dipping needle" makes with the horizon, when placed in the magnetic meridian, is called *the angle of dip*. This angle varies from place to place. Lines on a map connecting points of equal dip or inclination are called *isoclinic lines.* The angle of dip increases in the northern hemisphere, as we pass from the equator towards the north, and becomes 90° on the peninsula of Boothia Felix. Here the lines of induction enter the earth vertically. This point is called the *magnetic pole in the northern hemisphere.*

Since direction of the earth's field is not parallel to the horizon, a distinction must be made between the *horizontal* and the *vertical components* of the earth's magnetic intensity. These are denoted respectively by H and V, and are connected by the equation

$$\tan\theta = \frac{V}{H} \tag{305}$$

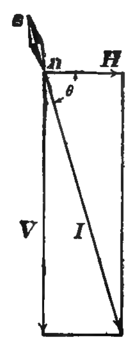


FIG. 127.

where θ is called the angle of dip or of inclination (Fig. 127). It represents the angle between the direction of the total intensity and its horizontal component. The magnitude of the total intensity of the earth's field is given by the equation

$$I = \sqrt{H^2 + V^2} \tag{306}$$

In Ann Arbor[1] $H = 0.19$ gauss

$V = 0.595$ gauss

$\theta = 73°$

In all civilized countries a systematic study of terrestrial magnetism is carried on by the government. Thus, the Coast

[1] *For the experimental determination of the horizontal component of the earth's magnetic field, see Manual, Exercises 76 and 77.*

and Geodetic Survey has established a number of permanent stations for this purpose, and also determines, from time to time, the magnetic elements in a large number of places uniformly distributed over the United States. Lately magnetic measurements have also been undertaken on the ocean by the department of terrestrial magnetism of the Carnegie Institution.

* **250. Secular Variations.** The three elements of terrestrial magnetism,—declination, inclination and intensity,—at a given place change in course of time. These variations consist either of slow, regular movements of the magnetic needle or of sudden and irregular disturbances.

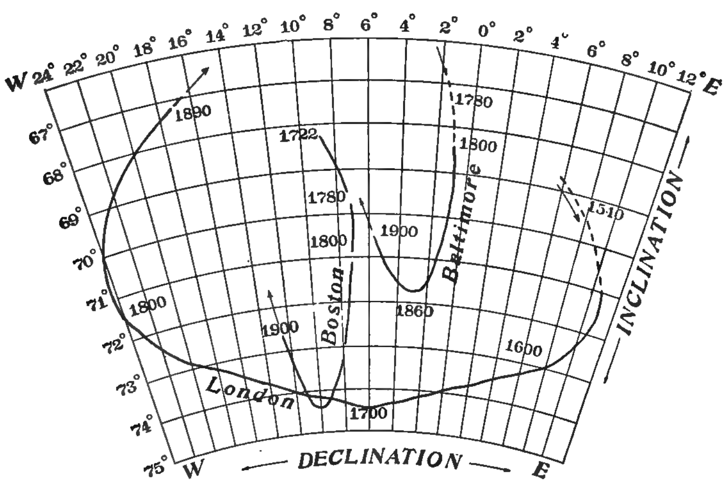


FIG. 129.

The most important of these is the slow, progressive change observed in the course of centuries. For example, the declination for London was 11° east in 1580; it diminished to zero in 1658, and then became west by an ever increasing amount until 1812, when it was 24° west. After that time it again decreased, and is now about 16° west, or still 27° from the value which it had in 1580. Similar slow changes are observed at all magnetic stations. In Boston the declination has changed from 7° west in 1800 to 13° west at the present time.

The inclination also shows secular variations. In Fig. 129 both variation of declination and inclination are plotted for London, Boston and Baltimore. The maximum inclination in this country was reached in the year 1860. The form of the curves suggests that the secular variation is a cyclic change, and that the curve will become a closed curve after a sufficiently long period of time. It is evident that the point called the magnetic pole of the earth is by no means a fixed point in the earth.

*** 251. Other Variations.** Besides the secular variations above noted, the following changes have also been observed:

(*a*) *Diurnal variations.* During the day the magnetic needle shows slight changes in its position, reaching the extremes between 8 and 10 A.M., and 1 and 3 P.M. The needle shifts slowly from the east towards the west during the morning, and returns in the opposite direction during the remainder of the day. The maximum variation from the mean declination, due to this cause, amounts to but a few minutes of arc.

(*b*) *Annual variations.* If the monthly values of the magnetic declination be corrected for the progressive secular change throughout the year, they exhibit a cyclic annual change, but this is only a fraction of a minute of arc and may, therefore, be neglected for all practical purposes. Similar minute variations depending upon the position of the moon with reference to the sun and earth have been detected, but they are even smaller than the annual variations.

(*c*) *Magnetic storms.* These are irregular disturbances which affect the magnetic elements and occur practically at the same time over large areas or in some cases affect the whole earth, progressing, according to Bauer, with a speed of about 7000 miles per minute. In exceptionally violent cases the needle may be deflected for a short time several degrees from its mean position. Small, spasmodic fluctuations of this kind occur frequently. The larger ones are often accompanied by auroral displays.

Magnetic disturbances of this kind seem to be more frequent and violent in years of maximum solar activity, as indicated by sun spots. Very little, however, is known concerning their

causes. In most cases where the disturbances extend over the whole earth they are probably produced by phenomena having their origin outside our earth.[1]

Problems

1. A magnetic pole of 15 C. G. S. units acts with a force of 4 dynes upon another pole at a distance in air of 6 cm. Find the strength of the second pole. *Ans.* 9.6 C. G. S. units.

2. Two equal bar magnets, each of pole strength 50 C. G. S. units and distance between the poles of 15 cm in air, are placed parallel to each other, 10 cm apart, both centers lying on the same perpendicular to the axes. The magnets point in opposite directions. Find the magnitude and direction of the force of attraction between the two magnets.

Ans. 41.5 dynes, at right angles to the magnets.

3. A needle having a magnetic moment of 12 C. G. S. units is placed in a uniform magnetic field, of intensity 16 gausses, in such a direction that it makes an angle of 30° with the lines of induction. Find the moment of the couple acting on the needle. *Ans.* 96 dyne-centimeters.

4. Calculate the intensity of the magnetic field at a point on the axis of a bar magnet 50 cm in air from its middle point, the strength of the poles being 100 units, and the distance between the poles 20 cm. *Ans.* 0.0347 gauss.

5. Find the direction and magnitude of the field intensity at a point 10 cm from the middle of the magnet of problem 4, the distance being measured in a direction at right angles to the magnet. Draw diagram.

Ans. 0.707 gauss, parallel to magnet, toward the south pole.

6. Show that the field intensity produced by a short magnet, at a point on its axis produced and at a considerable distance from it, is approximately $\frac{2M}{D^3}$, M being the magnetic moment and D the distance of the point from the middle point of the magnet.

7. What forces must be applied to a magnet whose magnetic moment is 500 C. G. S. units, in order to hold it in an east and west position in air, if the distance between the poles be 25 cm, and H be 0.19 gauss.

Ans. A force of 3.8 dynes on each pole.

8. A magnet is placed with its axis in the magnetic meridian and its south pole pointing north. It is found that there is a neutral point at a distance of 14 cm north from the south pole of the magnet. The distance between the poles is 10 cm, and H is 0.19 gauss. Find the pole strength of the magnet. *Ans.* 56.448 C. G. S. units.

[1] For a discussion of the causes of these disturbances, see Bauer, *Science*, vol. 33, p. 41, 1911.

ELECTRODYNAMICS

CHAPTER XXXIII

FUNDAMENTAL ELECTRICAL UNITS

252. Energy of Chemical Reaction. If a strip of zinc be placed in dilute sulphuric acid, the zinc is dissolved in the acid, forming zinc sulphate, while at the same time hydrogen gas appears upon the surface of the zinc. The displacement of the hydrogen of the acid by the zinc is a chemical reaction, and is accompanied by the evolution of a quantity of heat proportional to the amount of zinc dissolved. Energy is therefore liberated by this reaction, or the energy of the original substances is larger than the energy of the substances formed by the chemical reaction.

253. Simple Voltaic Cell. If a copper plate be placed in the same solution with the zinc, no change in the above-mentioned reaction can be observed so long as the two plates are not in contact. But as soon as the two plates are connected by a wire (Fig. 130), the process is quite different. Hydrogen bubbles now appear upon the copper plate, and if the heat produced be carefully measured, it will be found that the amount of heat appearing *in the liquid*, due to the solution of a given mass of zinc, is much smaller than in the first case. What has become of the remainder of the energy, set free by the chemical reaction?

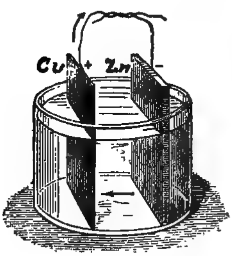


FIG. 130.

A careful measurement will show that the temperature of the wire connecting the two plates is higher than it was before, or that energy now appears *in the wire* in the form of heat.

The conclusion is evident that the two different metals in the dilute solution of sulphuric acid, when connected by a wire, afford a means of transforming the energy of chemical reaction into some other form of energy, which, in its turn, is transformed by the wire into heat. This new form of energy, which is quite distinct from any which has been studied thus far, is called *electrical energy.*

Any device which transforms the energy of chemical reaction into electrical energy is called a *voltaic* or an *electric cell.* The metal plates, to which the wires are connected, are called the *terminals, electrodes* or *poles* of the cell. The liquid joining the two plates is called the *electrolyte.* The system made up of cell and wire is called an *electric circuit.*

The simple arrangement described above was first used by Alessandro Volta[1] (1745–1827), and has therefore received the name, *the simple voltaic cell.* Many other types of cells have been devised since Volta's time, some of which are much more efficient than the original cell. A number of cells joined together is termed an *electric battery.* We shall return to the study of chemical generators of electricity in a subsequent chapter.

254. Magnetic Effect of an Electric Current. If we stretch a copper wire over a magnetic needle, parallel to its axis, and connect the ends of the wire to the terminals of an electric battery (Fig. 131), the needle will immediately be deflected and tend to place itself at right angles to the wire. This proves that electric energy reveals itself not only by heating the wire, but also by establishing a magnetic field about the wire. This important discovery was made by Oersted[2] in 1820.

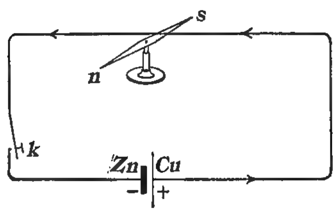


FIG. 131.

If the connections between the ends of the wire and the terminals of the battery be interchanged, the direction of the deflec-

[1] Volta, *Phil. Trans.*, 1800, p. 402.
[2] Oersted, *Gilbert's Ann.*, 66, p. 295, 1820

tion of the magnetic needle will be reversed. This experiment shows that the physical process going on in the wire has a *definite sense*, inasmuch as it may be considered as being either *positive* or *negative*, according to the direction in which the needle is deflected. It suggests that the wire serves as a carrier of a current of some kind, which produces both the magnetic and the heat effects, and that the current is reversed when the connections of the wire to the battery are reversed. The substance which may be assumed to flow through the wire, however, is certainly *not a material substance*, since the mass of the wire is not increased during the phenomenon.

An electric current, produced by a voltaic cell, may be defined as the *immaterial agent by means of which energy, set free by chemical reaction, is transferred from the cell to other parts of the circuit or to the space surrounding it.* Since electric currents flow readily through metals, metals are called *electrical conductors.*

This concept of electricity flowing in a circuit will be found to be very useful in the discussion and explanation of many phenomena to be studied in subsequent chapters.

255. Direction of an Electric Current. We have seen in the last paragraph that we must distinguish between a positive and a negative current, but we are at liberty to choose either direction in the wire as positive. The historical development of the subject has led to the general agreement to call the copper plate of a simple voltaic cell *the positive electrode* and the zinc plate the *negative electrode.* Consequently (Fig. 131), the positive current flows *from the copper through the wire to the zinc.* A current flowing in the *opposite sense* would be called a *negative* current.

It should be noted that while an electric current has both *magnitude* and *sense*, it has no definite *direction in space*, and cannot, therefore, be classed as a vector quantity.

256. Magnetic Field about a Current. The existence of a magnetic field about a wire carrying a current may be clearly shown by passing the wire through a sheet of cardboard whose

plane is perpendicular to the current. Iron filings sprinkled upon the cardboard will arrange themselves in concentric circles around the wire when the paper is gently tapped (Fig. 132). This proves that the lines of magnetic induction produced by a current are closed curves surrounding the conductor, and lie in a plane normal to the current.

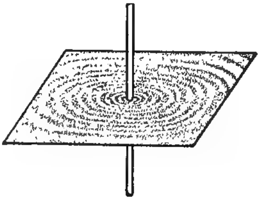


FIG. 132.

The direction and sense of this magnetic field at any point are shown by the position assumed by a short magnetic needle placed at that point. The tangential position of the needle gives the *direction*, and the position of the poles gives the *sense* of the field, since, if the current be reversed, the needle swings round through 180°, but still remains tangent to the circular lines of induction. This shows that the *sense* of the field is reversed with reversal of the current, while the *direction* of the field remains the same.

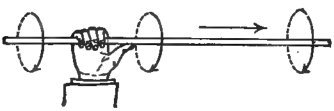


FIG. 133.

These experimental results may be expressed by the following rule:

Grasp the wire with the right hand, the outstretched thumb pointing in the direction of the current; then the fingers indicate the sense of the lines of magnetic induction (Fig. 133).

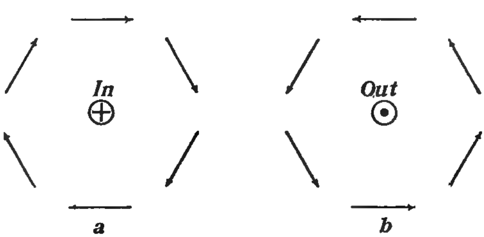


FIG. 134.

This relation may also be expressed by conceiving the current to be flowing *into the paper* through a section of the wire (Fig. 134 *a*), then the *sense* of the lines of induction is given by the *arrow heads;* if, on the other hand, the current be supposed to be flowing *out from the paper* through the section

of the wire (Fig. 134 *b*), *the reversed sense of the field* is shown by the *reversed arrow heads.* Or if the current be flowing toward the point of an auger, then the sense of the field is that in which the auger is turned to bore into the wood.

If a conductor carrying a current be bent into a loop, this loop will be found to have magnetic properties. From the preceding rule it can easily be seen that each part of the conductor contributes a number of lines of induction, all of them passing through the loop in the same direction (Fig. 135). The side of the loop where the lines enter has the properties of a magnetic *south pole;* the other side, those of a *north pole.* If we look towards a loop in which the current flows *counterclockwise,* the side of the loop facing us may be considered *a magnetic north-seeking pole;* if the current flow clockwise through the loop, the side facing us becomes a *south-seeking pole.*

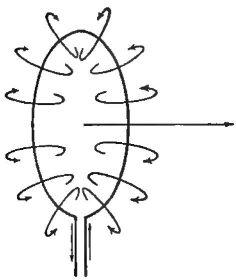


FIG. 135.

257. Magnetic Field due to a Circular Current. In order to measure a current by its magnetic effect, it is necessary to find some quantitative relation between the two. The simplest assumption is that the intensity of the magnetic field at any point P, in the neighborhood of a given conductor, is proportional to the current flowing through the conductor. Further, the intensity of the field at the point is the sum of the effects of all the elements of current, which may be obtained by dividing the conductor into very short sections.

From the experimental results obtained by Biot and Savart, upon the magnetic field due to a straight current, Laplace deduced the following empirical equation for the field intensity produced by each of these current elements:

$$H' = k \frac{Ids}{d^2} \sin \alpha \qquad (307)$$

where H' is the fraction of the field intensity contributed by

the element of the current I of length ds (Fig. 136), d the distance PA from the point to the current element, and α the angle between the direction of the current and the line PA.

The above empirical expression cannot be proven by direct experiment, since we can never deal directly with current elements, as here imagined, but the effect of conductors of definite length may be calculated from this law, usually by the use of calculus. Experimental results have fully proven the correctness of Laplace's law.

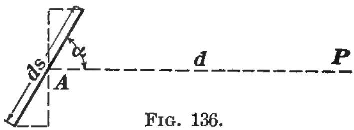


FIG. 136.

In the case of a circular current the total intensity of the field at the center is easily calculated, for in this case α is 90°, and d is equal to the radius of the circle for all elements, while the sum of all the elements Σds is the circumference of the circle, or $2\pi r$.

Thus we have in this case

$$H = k\frac{I}{r^2}\Sigma ds = k\frac{2\pi}{r}I \tag{308}$$

If the proportionality factor k be taken as unity, we have finally

$$H = \frac{2\pi}{r}I \tag{309}$$

258. Electromagnetic Unit of Current. In choosing the magnetic effect as a measure of the current, the proportionality factor k in equation (308) and the radius of the circle have been made equal to unity.

Hence, *unit current is that current which, when passing through an arc of unit length in a circle of unit radius will produce at the center of the circle a magnetic field of unit intensity.*

A unit pole placed at the *center of a circular coil of unit radius* through which unit current is flowing is acted upon by a mechanical force equal to 2π dynes.

The unit thus chosen is evidently based upon the units of the C. G. S. system of measurement and electromagnetic relations,

and is therefore called the C. G. S. *unit of current in the electromagnetic system.*

This unit is, however, found to be too large for practical purposes, and therefore one tenth of this value has been taken as the *practical unit of current,* and is called an *ampere,* after the French physicist, Ampère (1775–1836).

We shall see later how a current may be measured by the chemical effects which it produces in a solution of silver nitrate (Art. 285).

*** 259. The Tangent Galvanometer.** A galvanometer is an instrument in which the magnetic effect of a current is used either to detect or to measure small currents flowing through the galvanometer circuit.

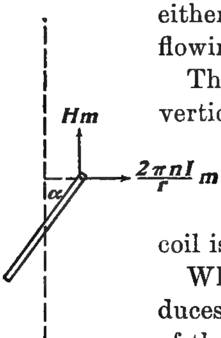


FIG. 137.

The tangent galvanometer consists of a large vertical circular coil, in the center of which is suspended a short magnetic needle whose length must be small in comparison with the radius of the coil. The plane of the coil is placed in the magnetic meridian.

When a current is sent through the coil, it produces a magnetic field perpendicular to the plane of the coil. Let n be the number of turns of wire in the coil, and r their mean radius; then the intensity of the magnetic field produced at the center of the coil by the current I (in C. G. S. units) is

$$H' = \frac{2\pi n}{r} I \tag{310}$$

Under the action of this field the needle will be deflected from the magnetic meridian through an angle α (Fig. 137).

If we call m the pole strength of the needle, l the distance between the poles, H the horizontal component of the earth's magnetic intensity, the deflecting moment due to the current is

$$\mathcal{T}' = H' \cdot ml \cos\alpha = \frac{2\pi nI}{r} ml \cos\alpha \tag{311}$$

and the restoring moment (Art. 237) is

$$\mathcal{T} = Hml \sin \alpha \tag{312}$$

Since for equilibrium these moments must be equal, we have

$$\frac{2\pi nI}{r} ml \cos \alpha = Hml \sin \alpha \tag{313}$$

or

$$I = \frac{r}{2\pi n} \cdot H \tan \alpha = \frac{H}{G} \cdot \tan \alpha \text{ C.G.S. units} \tag{314}$$

$$= 10 \frac{H}{G} \cdot \tan \alpha \text{ amperes} \tag{315}$$

Since $G = \frac{2\pi n}{r}$ is a constant for a given galvanometer, and H varies only by negligible amounts (Art. 251), the current is proportional to the tangent of the deflection produced by it, or

$$I = A \tan \alpha \tag{316}$$

If G and H be known by previous measurements, the tangent galvanometer may be used for measuring a current directly in C.G.S. units or in amperes.

260. The Movable Needle Galvanometer. The magnetic effect of a tangent galvanometer is usually quite small on account of the great distance of the coil from the needle and the small number of turns. In order to increase the effect, many turns of wire may be placed close to the needle. Of course the exact proportionality between the current and the tangent of the deflection is sacrificed by such an arrangement.

Frequently the effect is still further increased by the use of what is termed an *astatic pair of needles.* This consists of a pair of needles magnetized in opposite directions and connected by a thin rod. One of the needles has a slightly greater pole strength than the other. By this arrangement the turning moment, due to the earth's field, is made considerably smaller than it would be with a single needle. Galvanometer coils surround each needle, but are so wound that the turning moments exerted by the two coils upon the needle are in the

same sense (Fig. 138). Thus the effect of the current will be larger than it would be with a single coil and needle. Galvanometers built on this plan are extremely sensitive. The system of needles carries a small mirror, and the deflection of the mirror is determined by viewing through a telescope the image of a scale placed a short distance above or below the telescope and at right angles to it. One objection to the use of these galvanometers in ordinary work is found in the large influence which changes in the magnetic field about the galvanometer produce in the *zero point* or position of rest of the moving system. Such disturbances occur frequently in the neighborhood of conductors carrying a current, and can scarcely be avoided in a physical laboratory.

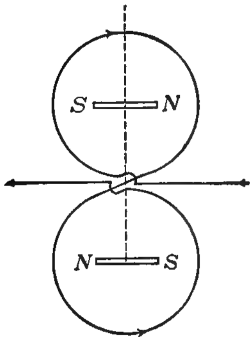


FIG. 138.

261. The D'Arsonval Galvanometer. The most satisfactory type of galvanometer for general use is the D'Arsonval galvanometer, which consists of a stationary magnet and a movable coil. The coil (Fig. 139 *a*) is suspended by means of a fine metal wire or ribbon, and is attached to the base of the instrument by another wire or metallic spiral spring. The current enters and leaves the coil by these upper and lower suspensions. The current flowing through the coil sets up a magnetic field (Art. 256), and the coil tends to place itself so that its lines of induction are parallel to those of the field of the stationary magnet. Thus with the current flowing as indicated, the side *ab* tends to rise up out of the paper. The turning moment is proportional to the current, but this moment is opposed

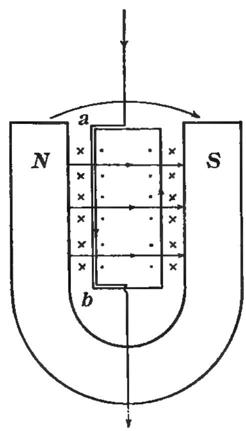


FIG. 139 *a*.

by the torque in the suspension, which tends to restore the coil to its original position. The coil comes to rest when these deflecting and restoring moments are equal. The finer the suspension, the more sensitive is the instrument. Deflections are usually observed by means of mirror and scale.

The great advantage of the D'Arsonval galvanometer (Fig. 139 *b*) is that its strong field, due to the permanent magnet, renders it entirely independent of the earth's magnetism, so that its readings are not at all affected by variations in the surrounding magnetic field. Additional advantages are found in the fact that it may be placed in any desired position, regardless of the magnetic meridian, and also that the movable system may easily be brought to rest by short-circuiting the swinging coil whereby its energy of swing is transformed into the energy of a small current (Art. 344).

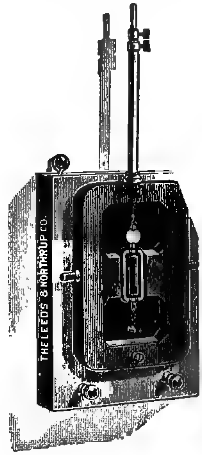


FIG. 139 *b*.

262. The Ammeter. Galvanometers, provided with a pointer and a scale so graduated as to indicate the current directly in amperes, are called *ammeters*. These instruments are usually of the D'Arsonval type. Fig. 140 represents an ammeter of the well-known Weston type. A *milammeter* is an ammeter of greater sensitiveness, reading to thousandths of an ampere. Any galvanometer may be used as an ammeter, provided it has been calibrated so that the exact relation between the deflections and the current passing through the instrument is known.

FIG. 140.

263. Quantity of Electricity. We have considered an electric current as a flow of electricity through a conductor. For a

constant current the quantity of electricity Q passing in a given time t through a circuit is proportional to the current and to the time, or

$$Q = It \tag{317}$$

Quantity of electricity is therefore measured by the product of a current into the time during which it flows. Conversely, we may say that a current is the *time rate of transfer of quantity of electricity.* The C. G. S. unit of quantity of electricity is the quantity transferred by a C. G. S. unit of current in one second.

The *practical unit of quantity* of electricity is called the *coulomb*, after the French physicist, Coulomb (1736–1806). It is the quantity of electricity transferred in *one second* by a current of *one ampere.* It is equal to 10^{-1} C. G. S. unit of quantity.

264. Resistance. It has been mentioned (Art. 253) that a conductor carrying an electric current is heated. Joule measured the amount of heat produced in a conductor by currents of different strengths, and discovered the law known as *Joule's law:*[1] *The heat produced by a current is proportional to the square of the current and to the time during which it flows.* In a mathematical form the law is written

$$H = RI^2t \tag{318}$$

The proportionality factor R, *which is constant for a given conductor*, is called the *electric resistance. Electric resistance is therefore a characteristic property of a conductor, by virtue of which the energy of an electric current is transformed into heat.*

In defining the unit of resistance, heat should not be expressed in calories, but in ergs or joules (Art. 177); consequently the C. G. S. unit of resistance is that resistance in which a quantity of heat equal to *one erg* is produced *in one second by a* C. G. S. *unit of current.*

This unit is much too small for practical purposes, and it has therefore been agreed to take the unit of resistance 10^9 times as large. This unit is called the *ohm*, after the German physicist, Ohm (1789–1854).

[1] Joule, *Phil. Mag.* 18, p. 308, 1841.

In accordance with the resolutions of the International Electrical Conference of London, 1908,[1] the *ohm*, the *ampere* and the *volt* (Art. 267), defined in terms of the C. G. S. units, were made the *fundamental electrical units*. These, however, were recognized as purely ideal units. As a system of concrete, practical units, representing these ideals, "and sufficiently near to them for purposes of electrical measurements and as a basis for legislation," the *international ohm*, the *international ampere* and the *international volt* were defined and their adoption was recommended. Specific definitions of these international units will be given in their appropriate places.

Accurate measurement of resistance involves the comparison of the resistance of a conductor with that of a concrete standard. Such standards have been prepared by the National Physical Laboratories of various countries in accordance with the definition of the international ohm.

"*The international ohm is the resistance offered to an unvarying electric current by a column of mercury at the temperature of melting ice, 14.4521 grams in mass, of a constant cross-sectional area, and of a length of 106.300 centimeters.*"

Since this concrete unit agrees with the ohm within the degree of accuracy attainable by the most refined methods of measurement of the present time, we shall make no further distinction between these units, but shall use simply the term *ohm*.

265. Difference of Potential. In the last article we have seen that electric energy may be expressed as

$$Energy = I^2Rt$$

or if quantity of electricity Q be substituted for It, we may write

$$Energy = IR \cdot Q \tag{319}$$

We may think of the heat generated in the conductor as the equivalent of the work done by electrical agencies in order to force a quantity of electricity Q through the conductor. From

[1] *London Electrician*, vol. 62, p. 104, 1908.

this point of view the product IR in (319) measures *the work done in carrying unit quantity of electricity through a resistance R.*

The mechanical analogy between an electric current and a liquid under pressure, flowing through a system of pipes which offer some resistance to the flow, leads naturally to the concept of a difference of electric pressure which must exist between the ends of a conductor carrying a current. We call this quantity *electric pressure*, or better, *difference of electrical potential.* Without admitting the necessity for such a mechanical picture, we may state, however, that in a resistance R through which a constant current I is flowing, there exists always a difference of potential, measured by IR, or

$$V_1 - V_2 = IR \tag{320}$$

From this point of view V_1 and V_2 may be called the *potentials* at the terminals of the resistance, though nothing be known about their absolute value. It is further to be noted that in general the current flows from points of higher to points of lower potential, just as water flows from points of higher to points of lower level.

266. Electromotive Force. The mechanical analogy, mentioned in the last article, led, in the early development of the theory of electricity, to the assumption that every generator of electricity, such as an electric cell, produces the difference of electric pressure which causes the current to flow. From this point of view we may compare an electric cell to a pump, which continuously drives the electric fluid through the system of conductors. This theory brought into general use the misleading term *electromotive force*, often written E. M. F.

Electromotive force is a difference of potential, considered as the cause of an electric flow through a resistance, thereby producing a continuous fall of potential throughout the whole circuit. Electromotive forces are usually localized in definite portions of the circuit, and may be recognized by the appearance at these places of an *increase* of potential, as we pass along the circuit in the direction in which the current flows, or would

flow if the circuit were closed. Thus if the terminals of a simple voltaic cell be joined by a conductor, there is a continuous drop of potential along the conductor from the copper to the zinc, but a sudden rise from the zinc to the copper as we pass through the cell back to the copper electrode. The cell is, therefore, the seat of an electromotive force, and this is always present, even on an open circuit.

In that part of a circuit which does not contain an E. M. F. there will be no difference of potential, unless a current flow through it (eq. 320). These differences of potential over the various parts of a closed circuit may be considered as the result of the E. M. F. which causes the current to flow, and their sum $I\Sigma R$, taken over the whole circuit, measures the E. M. F. Electromotive force is therefore the work per unit quantity of electricity spent in the *whole circuit.* The difference of potential over a part of a closed circuit is always smaller than the electromotive force present in the whole circuit.

267. Unit Difference of Potential. Unit difference of potential, or unit electromotive force, is the difference of potential produced at the terminals of unit resistance when traversed by unit current.

The C. G. S. unit is too small for practical purposes. The *practical unit of difference of potential* is 10^8 C. G. S. units. It is that difference of potential *which, when steadily applied to a conductor whose resistance is one ohm, will produce a current of one ampere.* It is called the volt, after the Italian physicist, Volta (1745–1827).

268. Voltmeters. Any instrument designed to measure difference of potential is called a voltmeter. Suppose it be desired to measure the difference of potential between two points A and B of a circuit (Fig. 141). If we connect a conductor $ADCB$ to these points, the difference of potential will produce a current through this conductor, according to equation (320), and the

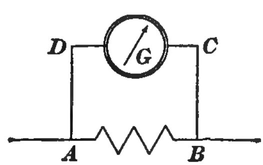


FIG. 141.

difference of potential measured over this circuit will be the same as over AB, and equal to the original difference of potential, *unless the current through the original circuit is appreciably changed by the introduction of the new resistance.* If $ADCB$ form the coil of a galvanometer G, the current through this instrument produces a deflection. Evidently the galvanometer may be calibrated in such a manner that the readings give directly the difference of potential existing at its terminals. The readings of such an instrument therefore indicate, not the current, but the product of the current flowing through it into its *own resistance.* The general form of the voltmeter (Fig. 142) is the same as that of the ammeter.

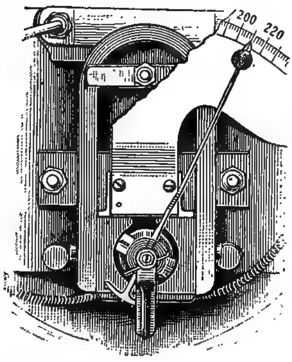


FIG. 142.

269. Electric Energy, Electric Power. Electric energy is energy measured in terms of electric quantities. We have seen (Art. 264) that it may be expressed as

$$Electric\ Energy = I^2Rt$$

or, introducing the difference of potential E,

$$Electric\ Energy = EIt = EQ \tag{321}$$

Remembering that an ampere is 10^{-1} C. G. S. unit, an ohm 10^9 C. G. S. units and a volt 10^8 C. G. S. units, electrical energy is given in joules or 10^7 ergs, if the electric quantities be measured in practical units. A *volt coulomb* is therefore identical with *one joule.*

Electric power is the time rate of expenditure of electrical energy, or

$$Electric\ Power = EI \tag{322}$$

One *volt ampere* is identical with one *watt.* A *kilowatt* is 1000 watts. In commerce, the unit of energy used is the *watt hour* = 3600 joules, or the larger unit, the *kilowatt hour* = 1000 watt hours.

If it be desired to measure the heat in calories, produced by the absorption of electrical energy, we have, since one joule equals 0.24 calorie (Art. 177),

$$H = I^2Rt \text{ joules} = 0.24\, I^2Rt \text{ calories} \tag{323}$$

where all electric quantities are measured in practical units.

CHAPTER XXXIV

OHM'S LAW AND ITS APPLICATIONS

270. Ohm's Law. Ohm [1] found in 1827 that the resistance of a given conductor is independent of the magnitude and the direction of the current flowing through it. Equation (320), in which we have made this assumption, is therefore called Ohm's law. It is frequently written in the equivalent form

$$I = \frac{V_1 - V_2}{R}$$

The current flowing through a conductor is proportional to the difference of potential and inversely proportional to the resistance of the conductor. This law holds for all *constant* currents or currents whose strength changes very little in course of time.

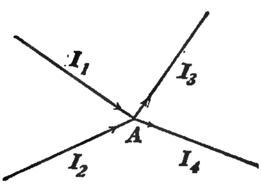


FIG. 143.

271. Kirchhoff's Laws.[2]

First Law. If several conductors meet at a point, the algebraic sum of all the currents flowing toward the point is zero. A current flowing from the point must be taken as negative. Thus, in (Fig. 143),

$$I_1 + I_2 - I_3 - I_4 = 0 \tag{324}$$

or, in general

$$\Sigma I = 0 \tag{325}$$

Kirchhoff's first law simply states that no electricity accumulates at any point of a closed electric circuit. In a simple circuit the current is the same, no matter in what part of the circuit the current is measured.

[1] Ohm, *Die Galvanische Kette*, 1827.
[2] Kirchhoff, *Pogg. Ann.* 72, p. 497, 1847.

Second Law. In any closed circuit the sum of the products of the resistance of each part into the current flowing through it is equal to the sum of the electromotive forces in the circuit. Of course, each current and each electromotive force must be taken with the proper sign.

Kirchhoff's second law may be written in the form

$$\Sigma IR = \Sigma E \tag{326}$$

This law is an extension of Ohm's law, for the current through a simple circuit containing an E.M.F. E and total resistance R is

$$I = \frac{E}{R} \tag{327}$$

Such a circuit, made up of a cell, of electromotive force E and conductors joining the terminals of the cell, may be considered as consisting of two parts: (*a*) the *external circuit*, of resistance R; and (*b*) the *cell*, whose resistance r between its terminals is called the *internal resistance*. In accordance with the second law, we have

$$E = IR + Ir = I(R + r) \tag{328}$$

272. Wheatstone's Bridge. Wheatstone's bridge is a device for the measurement of a resistance by comparing it with a known resistance. It consists of a network of six conductors, joining four points A, B, C and D (Fig. 144), so arranged that each point is joined to each of the other three points by separate conductors.

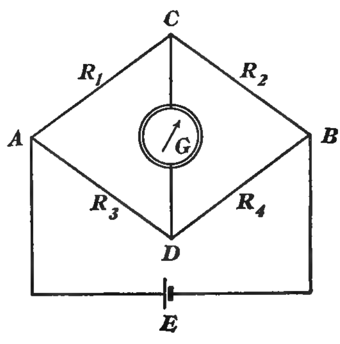


FIG. 144.

Let one of the conductors contain an E.M.F., for example, a cell E; four of the others will form a divided circuit, while the remaining one, containing a galvanometer G, will form a bridge between the two parallel conductors. Let R_1, R_2, R_3, R_4 be the resistances of the four

branches of the divided circuit, and suppose them to be so adjusted that no current flows through the galvanometer. Then it may be shown that the resistances satisfy the relation

$$\frac{R_1}{R_2} = \frac{R_3}{R_4} \tag{329}$$

For, let the current through AC be I_1 and the current through AD be I_2. Since there is no current through the galvanometer, the current through CB must be I_1, and that through DB, I_2. The difference of potential from A to C is I_1R_1, and that from A to D is I_2R_3. For the closed circuit $ACDA$, since there is no current between C and D, and no E. M. F. in this circuit, we have by Kirchhoff's second law

$$I_1R_1 - I_2R_3 = 0$$

or

$$I_1R_1 = I_2R_3 \tag{330}$$

Similarly, for the closed circuit $CBDC$ we have

$$I_1R_2 = I_2R_4 \tag{331}$$

Dividing (330) by (331) we obtain

$$\frac{R_1}{R_2} = \frac{R_3}{R_4}$$

or

$$R_1 = R_2\,\frac{R_3}{R_4} \tag{332}$$

From the above relation it is evident that if three of the resistances be known, the fourth may at once be determined. In fact, it is sufficient to know one resistance and *the ratio between the other two.*[1]

273. Laws of Resistance. Careful measurements of resistance have established the following experimental facts:

(1) The resistance of a conductor of uniform cross section is proportional to its length.

[1] *For the application of the Wheatstone bridge to the measurement of resistance, see Manual, Exercises 57–61.*

(2) The resistance of a conductor of uniform cross section is inversely proportional to its cross-sectional area.

(3) The resistance of a conductor depends upon the material of which it is made.

These three laws may be combined in the equation:

$$R = \rho \frac{l}{a} \tag{333}$$

274. Resistivity. The proportionality factor of equation (333) is a constant for a given substance and is called the resistivity, or specific resistance of the substance. It is easily calculated from the resistance R of a conductor, made of this material, and having a definite length l and uniform cross-section a, from the relation

$$\rho = \frac{Ra}{l}$$

It is numerically equal to the resistance between the opposite faces of a cube of the substance whose edge is one centimeter. The unit of resistivity is the ohm-centimeter.

TABLE XVII

RESISTIVITY OF VARIOUS SUBSTANCES, AT 18° C, IN OHM-CENTIMETERS

SUBSTANCE	RESISTIVITY	SUBSTANCE	RESISTIVITY
Aluminium	3.2×10^{-6}	Mercury	95.8×10^{-6}
Copper	1.7 "	Nickel	10 "
German silver	30 "	Platinum, pure	10.8 "
Iron	12 "	Platinum, commercial	14 "
Manganin	42 "	Silver	1.6 "

275. Conductance and Conductivity. The reciprocal of the resistance of a conductor is called its *conductance*, and the reciprocal of the resistivity of a substance is termed its *conductivity*.

It should be noted that the former refers to a characteristic property of a given conductor, the latter to a property of the substance of which it is made.

276. Resistances in Series. If a number of resistances R_1, R_2, R_3, $\cdots$ R_n be joined in series, the total resistance R is equal to the sum of the resistances of the separate conductors. For, let E be the difference of potential between the ends of the first and last conductor, and I the current in the circuit, then

$$E = IR$$

But by Kirchhoff's second law :

$$E = I(R_1 + R_2 + \cdots + R_n) \tag{334}$$

Therefore
$$R = R_1 + R_2 + \cdots R_n \tag{335}$$

277. Resistances in Parallel. If a number of resistances be joined in parallel (Fig. 145), and E denote the difference of potential at the ends A and B of the parallel system, R the total resistance between A and B, and I the total current, then

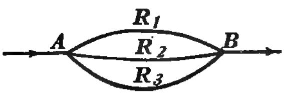


FIG. 145.

$$I = \frac{E}{R}$$

But for each separate conductor

$$I_1 = \frac{E}{R_1}, \qquad I_2 = \frac{E}{R_2}, \text{ etc.}$$

By Kirchhoff's first law

$$I = I_1 + I_2 + \cdots + I_n$$

or
$$\frac{E}{R} = \left(\frac{1}{R_1} + \frac{1}{R_2} + \cdots + \frac{1}{R_n}\right)E \tag{336}$$

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \cdots + \frac{1}{R_n} \tag{337}$$

The conductance of a system of parallel conductors, therefore, equals the sum of the conductances of the separate branches.

In the case of two conductors the last equation reduces to

$$R = \frac{R_1 R_2}{R_1 + R_2} \tag{338}$$

In the case of two resistances joined in parallel, either branch may be considered as a *shunt* to the other, and the currents through the branches are, of course, inversely proportional to the resistances of the branches.

*** 278. Change of Resistance with Temperature.** As a rule the resistance of metallic conductors increases with the temperature, and this change is nearly proportional to the temperature change, or

$$R_t = R_0(1 + \alpha t) \tag{339}$$

where α is called the *temperature coefficient of resistance.* For pure metals α is nearly 0.004 per degree centigrade.

The large temperature coefficient of pure metals renders them unsuitable for accurate standards of resistance. For this reason alloys, such as German silver, constantan and manganin are used in the construction of standards of resistance. These alloys have smaller temperature coefficients than pure metals. Manganin, consisting of 84 per cent copper, 12 per cent manganese and 4 per cent nickel, has, at ordinary temperatures, a very small positive coefficient, which at higher temperatures becomes zero, and finally negative.

The resistance of carbon and of all electrolytes decreases with rise of temperature.

279. Conductors and Insulators. We have seen that metals are conductors of electricity, or that they allow the passage of an electric current when a difference of potential is established between two points in the metal.

Some other substances, such as carbon, certain oxides and a few minerals also show metallic conduction. They are, however, relatively poor conductors, their conductivity being much smaller than that of metals.

Other substances offer very high resistance to the passage of an electric current, or even prevent its passage altogether. Such substances are called *insulators* or *dielectrics.* They are used extensively to prevent the passage of electricity from one conductor to another or to the earth, as, for example, from the trolley line of a street car or from telegraph or telephone lines.

The following substances are good insulators: Cold, dry glass, porcelain, ebonite, paraffine, sulphur, mica, fur, feathers, dry gases, etc. Whenever one of these substances is introduced into a metallic circuit containing an E.M.F., the current is interrupted and the circuit is said to be open.

Problems

1. At a place where $H = 0.19$ gauss, a certain current causes a deflection of 25° in passing through a tangent galvanometer, of which the coil, 30 cm in diameter, is made up of ten turns of wire. Determine the current in amperes. *Ans.* 0.211 ampere.

2. What is the constant A of a tangent galvanometer, in which a current of 5×10^{-4} ampere produces a deflection of 10°? How large is the constant G of the galvanometer, if H be 0.19 gauss?
Ans. (*a*) 0.002836.
(*b*) 670.1.

3. How large a quantity of electricity passes through the tangent galvanometer of problem 2 in one minute, if the deflection be kept constant at 45°? *Ans.* 0.17016 coulomb.

4. A 16 candle power incandescent lamp on a 110-volt circuit uses a current of 0.5 ampere. How much electric power is absorbed in one lamp? How much per candle? How much energy, expressed in calories, is absorbed by the lamp in one hour?
Ans. (*a*) 55 watts; (*b*) 3.4375 watts per candle; (*c*) 47300 calories.

5. In a house 10 incandescent lamps (see problem 4) are kept burning each evening for 3 hours. What will be the bill for 30 days, if the price per kilowatt-hour be 12.5 cents? *Ans.* $6.19.

6. If only 14 per cent of the energy spent in an incandescent lamp is transformed into light, how much heat, expressed in calories, is given off in an hour by a lamp of 220 ohms resistance and placed on a 110-volt circuit? *Ans.* 40,867 calories.

7. A copper calorimeter of mass 100 g, containing 500 g of water at 20° C, is heated by an electric current of 3 amperes, the resistance of the heating coil being 20 ohms. Compute the temperature of the calorimeter after 6 minutes. Neglect the increase of resistance, due to heating and the effect of radiation. *Ans.* 50°.53 C.

8. If electrical energy costs 12.5 cents a kilowatt-hour, what is the cost per calorie? *Ans.* 14.53×10^{-6} cent.

9. How much heat is generated in an electric iron, using 3.5 amperes from a 110-volt circuit for 1 hour? How much will it cost to use it an hour? *Ans.* (*a*) 332,640 cal.
(*b*) 4.8 cents.

10. The resistance ADB of a Wheatstone bridge (Fig. 144) consists of a uniform wire 100 cm long. The resistance R_2 is 65 ohms. The point D is 25 cm from A when no current passes through the galvanometer. Compute the resistance R_1. *Ans.* 21.66 ohms.

11. What must be the cross section of a wire, 150 cm long, in order to furnish the same resistance as a wire of the same material 80 cm long and of 1 mm^2 cross section? *Ans.* 1.875 mm^2.

12. What must be the length of an iron wire in order that its resistance may be the same as that of a copper wire 100 cm long and of twice the diameter of the iron wire? *Ans.* 3.542 cm.

13. What will be the relative weight of a copper and of an aluminium wire of equal length and of such cross sections that both wires have the same resistance? Specific gravity of copper = 8.9; of aluminium = 2.06.

Ans. As 2.295 to 1.

14. Compute the resistance of 100 meters of copper wire, 1 mm in diameter, at 18° C. Compute the resistance at the same temperature of an iron wire of the same length and diameter.

Ans. (*a*) 2.1645 ohms.
(*b*) 15.28 ohms.

15. An electric cell having an E. M. F. of 1.5 volts is connected by a copper wire 2 meters long and 0.5 mm in diameter. The internal resistance of the cell is 0.5 ohm. Compute the current. (Temperature = 18° C.) *Ans.* 2.228 amperes.

16. A wire having a resistance of 10 ohms carries an electric current. In order to measure the difference of potential at its terminals, a voltmeter of 300 ohms resistance is attached (Fig. 141) to the terminals, the total current remaining unchanged. The voltmeter reads 3 volts. What was the difference of potential before the voltmeter was attached? *Ans.* 3.1 volts.

17. The total current in a circuit consisting of three parallel wires is 1 ampere. The three wires are all of the same material and cross section, but of lengths 50, 30 and 10 cm. Compute the current through each wire. *Ans.* (*a*) 0.1304 ampere; (*b*) 0.2174 ampere; (*c*) 0.6522 ampere.

18. It is desired to reduce to one tenth the deflections of a galvanometer, of resistance *G*, by means of a parallel resistance or shunt. Compute the resistance of the shunt. *Ans.* $\frac{1}{9}$ G.

19. Find the resistance of a copper wire at 25° C, if its length be 150 cm and its cross section 0.75 mm^2. *Ans.* 0.03495 ohm.

20. A circuit consists of a cell of E. M. F. 1.5 volts and internal resistance 0.5 ohm, and a copper wire of 50 ohms resistance at 0° C. Compute the current at 30° C, assuming that neither the E. M. F. nor resistance of the cell change with temperature. *Ans.* 0.02655 ampere.

CHAPTER XXXV

ELECTROLYSIS

280. Electrolytes. Liquids differ greatly in their electrical behavior. Almost all organic liquids, such as paraffine oil, kerosene, etc., are insulators, and even pure water is a very poor conductor of electricity. Certain other liquids, however, notably aqueous solutions of acids, bases and salts are good conductors. To this latter class belong also solutions in liquid ammonia, in formic acid and in some alcohols. Since these solvents are of little practical importance in comparison with water, we shall restrict ourselves to the study of aqueous solutions.

If two platinum strips which are connected to a source of electricity be dipped into a dilute solution of sulphuric acid, a current passes through the liquid and at the same time gas bubbles are seen to rise from the platinum strips, showing that the electric flow is accompanied by a chemical decomposition of the liquid.

Conductors which undergo chemical decomposition when traversed by an electric current are called *electrolytes.*

A vessel containing an electrolyte and supplied with solid conductors dipping into the electrolyte is called an *electrolytic cell.* The solid conductors are called *electrodes.* The electrode through which the current enters is the *anode*, and the one through which it leaves the electrolyte is called the *cathode.*

281. Electrolysis of Sulphuric Acid. Invert over the platinum electrodes of an electrolytic cell (Fig. 146), containing dilute sulphuric acid, two graduated test tubes completely filled with the electrolyte. When a current is sent through the solution,

bubbles of gas rise from the electrodes and collect in the upper parts of the tubes. The volume of the gas above the cathode is twice that above the anode. The former gas can be proven to be hydrogen by the slight explosion which ensues when it is tested with a burning match. The gas collected above the anode will relight the glowing end of a match, thus showing that it is oxygen.

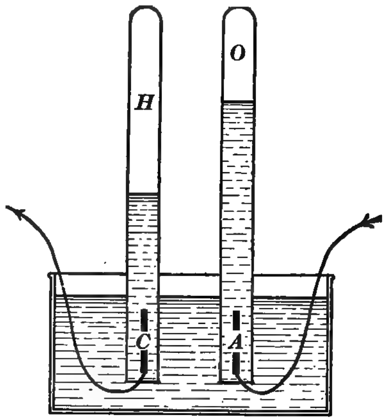


FIG. 146.

Since the gases appearing at the electrodes are always found in the ratio of *two volumes of hydrogen* to *one of oxygen*, exactly in accordance with the chemical symbol for water (H_2O), it would appear at first glance that the electrical action has been primarily the electrolysis of water, without affecting the acid in any way. However, experimental evidence leads to the conclusion that the real conductor of the electric current is the acid in the solution rather than the water. Thus, if the above experiment be repeated with zinc instead of platinum electrodes, hydrogen is produced at the cathode as before, but no oxygen appears at the anode. On the contrary, zinc goes into solution, forming zinc sulphate ($ZnSO_4$). We must therefore conclude that when the sulphuric acid (H_2SO_4) is decomposed, the hydrogen appears at the cathode as before, but in this case the acid radical (SO_4) unites with the zinc anode, forming zinc sulphate. If the electrodes be insoluble, as in the case of platinum electrodes, the (SO_4) decomposes the water around the anode, reproducing H_2SO_4 and liberating oxygen.

282. Electrolysis of Metallic Salts.

(*a*) *Electrolysis of copper sulphate.* If the experiment of the last article be repeated with a solution of copper sulphate ($CuSO_4$), as electrolyte, between platinum electrodes, no hydro-

gen appears at the cathode, but the platinum cathode itself will soon be covered with a layer of metallic copper. The copper in this case plays the rôle of the hydrogen in the last experiment. At the anode oxygen is liberated as before.

(*b*) *Electrolysis of lead acetate.* If a solution of lead acetate be electrolyzed between platinum electrodes, oxygen will appear at the anode as before, while lead will be deposited upon the cathode in the form of shining, fern-like lead crystals. The resemblance of this growth to that of a tree is very striking indeed. Hence the term *lead tree.*

(*c*) *Electrolysis of sodium sulphate.* If a solution of sodium sulphate (Na_2SO_4) be electrolyzed between platinum electrodes, we shall see that oxygen and hydrogen are liberated at the electrodes as in the electrolysis of sulphuric acid. If, however, we add to the cell some sensitive indicator for acids and alkalies, we shall get a better insight into the real reaction which is taking place.

Purple cabbage, when steeped in warm water for a few hours, yields a deep purple fluid which turns red in the presence of an acid and green in the presence of an alkali. When this cabbage solution is added to a neutral solution of sodium sulphate, the liquid is of a uniformly dark blue or purplish color. On passing a current through this solution, the gases rise from the electrodes as before, but the liquid turns red about the anode and green about the cathode. This shows that, while hydrogen and oxygen are given off, the sodium has been liberated at the cathode and the SO_4 radical at the anode.

The free sodium at the cathode acts chemically upon the water of the electrolyte, forming the alkaline sodium hydroxide ($NaOH$), and liberating hydrogen. At the anode the free radical SO_4 unites with water, forms H_2SO_4 and liberates O. If the contents of the electrolytic cell be carefully poured out into a clean beaker, the solution at once assumes its dark blue color, thus showing that *equivalent amounts of alkali and acid* have been developed by the electrolysis, and that, when reunited, the solution is again neutral. It also shows that this secondary chemical reaction manifests itself *only at the electrodes while the*

electrolytic action takes place through the liquid *without affecting this sensitive indicator in the least.*

From these experiments we may conclude that in a salt solution the dissolved substance is decomposed by the passage of an electric current. *Hydrogen and all other metals are liberated at the cathode, while non-metals and acid radicals appear at the anode.* The theory of electrolytic dissociation which attempts to explain the mechanism of the conduction of electricity through electrolytes will be given later.

283. Faraday's Laws of Electrolysis.[1] The following fundamental laws of electrolysis were first established by Faraday:

1. *The mass of an electrolyte decomposed by an electric current is directly proportional to the quantity of electricity passing through it.*

2. *If the same quantity of electricity pass through different electrolytes, the masses of the substances liberated at the electrodes are proportional to their chemical equivalents.*

The chemical equivalent of a substance is its atomic mass divided by its valence. For example, the atomic mass of hydrogen is 1, that of oxygen 16, of silver 107.9 and of copper 63.6. The valence for hydrogen is 1, for oxygen 2, for silver 1 and for copper in copper sulphate 2. Therefore the same quantity of electricity which would liberate one gram of hydrogen liberates 8 grams of oxygen, 107.9 grams of silver from a solution of silver nitrate and 31.8 grams of copper from a copper sulphate solution.

284. Electrochemical Equivalent. If we denote by M the mass of a substance liberated by a constant electric current I, flowing for t seconds through an electrolytic cell, Faraday's first law may be written

$$M = zIt \tag{340}$$

where z is called the *electrochemical equivalent of the substance*, and may be defined as the *mass per coulomb liberated by electrolytic action.*

[1] Faraday, *Researches, 3d series*, paragraph 377. 7th series, paragraph 783. Dec. 31, 1833.

Faraday's second law may be written

$$\frac{z}{z_h}=\frac{\frac{m}{v}}{\frac{m_h}{v_h}}$$

where m and v denote the atomic masses and valences, and the subscript h refers to hydrogen.

Since both m_h and v_h are unity

$$z=\frac{m}{v}z_h \tag{341}$$

Faraday's laws may be combined to

$$M=\frac{m}{v}z_h It \tag{342}$$

The electrochemical equivalent of hydrogen has been found by experiment to be 0.00001036 *gram per coulomb*. Therefore for any other substance

$$z=0.00001036\frac{m}{v} \text{ grams per coulomb} \tag{343}$$

The quantity Q of electricity necessary to liberate the mass of one chemical equivalent is evidently

$$Q=\frac{1}{0.00001036}=96{,}530 \text{ coulombs} \tag{344}$$

285. Definition of the Ampere. Unit current was defined (Art. 258) in terms of the intensity of the magnetic field which it produces. It is, however, entirely feasible to measure a current by its chemical effect. We need to know only the exact electrochemical equivalent of a substance, the mass of this substance deposited upon an electrode, and the time during which the constant current flows. Equation (340) gives then directly:

$$I=\frac{M}{zt} \text{ amperes} \tag{345}$$

Instruments designed to measure a current by its electrochemical effect are called *coulometers* or *voltameters*. The first of the two terms is preferable, since the mass deposited depends

primarily upon the quantity of electricity, that is, upon the number of coulombs.

For work of moderate accuracy the copper coulometer, that is, an electrolytic cell with copper electrodes in an aqueous solution of copper sulphate, is used. The electrochemical equivalent of copper from this solution is 0.000329 gram per coulomb.[1]

For most accurate determinations of quantity of electricity, the silver coulometer is preferred. This is an electrolytic cell, in which a platinum bowl filled with a solution of silver nitrate forms the cathode, while the anode, a plate of electrolytic silver, is suspended in the solution (Fig. 147). When the current passes, silver is deposited upon the platinum bowl, and the mass of the deposit can be determined with great accuracy by means of the balance.

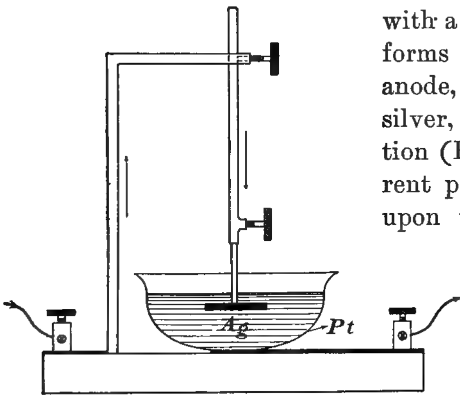


FIG. 147.

The *international ampere* is defined as "*the unvarying electric current, which, when passed through a solution of nitrate of silver in water . . . deposits silver at the rate of 0.00111800 of a gram per second.*" As remarked in Art. 264, the fundamental and international electrical units are so nearly identical as to justify no further discrimination between the ampere and the international ampere.

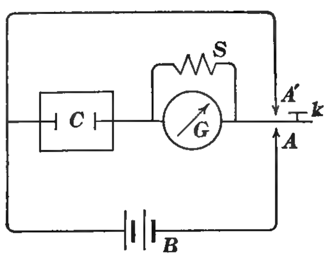


FIG. 148.

286. Polarization. Arrange an electric circuit as shown in Fig. 148. If the key k be connected

[1] *For the measurement of current by the coulometer, see Manual, Exercise 70.*

to the point A, the battery B sends a current through the shunted galvanometer G and the electrolytic cell C, consisting of platinum electrodes in sulphuric acid. It will be seen that the deflection of the galvanometer decreases. This decrease is due, in large measure, to the appearance of an E. M. F. in the electrolytic cell, which tends to send a current in the opposite direction through the circuit. The existence of this *counter E. M. F.* is shown by opening the key at A and connecting it to A'. A current now flows through the galvanometer in the opposite direction to the original current. To render this more evident the shunt S may be opened before making connection with A'.

This production of a counter E. M. F. by electrolytic action is called *polarization*. It is explained by the constant tendency of the products of electrolysis to reunite and to return into the solution.

Let E be the E. M. F. of the battery B, E' the counter E. M. F. due to polarization; then, according to Kirchhoff's second law,

$$I = \frac{E - E'}{R} \tag{346}$$

where R is the total resistance of the circuit.

The E. M. F. of polarization with mercury electrodes in sulphuric acid may reach a value of more than two volts. Electrolytic polarization occurs in all cases in which the metal in the electrolyzed salt is different from the material of the electrodes. If the two be chemically identical, little or no polarization is observed. For example, if the above experiment be repeated with two copper electrodes in a solution of a copper salt, copper goes into solution at the anode and is deposited at the cathode. No change in the electrodes and practically no polarization occurs upon closing the circuit. Electrodes of this kind are called unpolarizable. The only polarization possible is that due to a change of concentration of the electrolyte about the electrodes.

* **287. Electrolytic Resistance.** Electrolytes offer a resistance to the passage of an electric current. Joule's, Ohm's and

Kirchhoff's laws hold for the resistance of electrolytes. It is more usual, however, to speak of the *conductivity* of an electrolyte rather than of its resistance.

The conductivity of electrolytes is usually expressed either as *molecular conductivity* or as *equivalent conductivity*. Molecular conductivity is the conductivity of an electrolyte divided by its concentration. The concentration is usually expressed in terms of the number of grammolecules N, per liter of the solution. A grammolecule of a substance is a mass of the substance in grams equal to its molecular mass. The molecular concentration per cm^3 is therefore $\frac{N}{1000}$.

If the conductivity be called k and the molecular conductivity κ, then

$$\kappa = \frac{1000}{N}k$$

This is evidently numerically equal to the conductance, between electrodes one centimeter apart, of the volume of the solution containing one grammolecule of the dissolved substance.

The equivalent conductivity of an electrolyte is the conductivity divided by the concentration of gram equivalents of the dissolved substance, or its molecular conductivity divided by the valence v, of the metal in the compound. If n or $\frac{N}{v}$ be the number of gram equivalents in a liter of the solution, and k the conductivity, then the equivalent conductivity is

$$\lambda = \frac{1000}{n}k \tag{347}$$

The equivalent conductivity is therefore numerically equal to the conductance, between electrodes one centimeter apart, of so much of the solution as contains one gram equivalent of the dissolved substance.

It is found by experiment that the equivalent conductivity of most electrolytes increases with increasing dilution and reaches a limiting value for very dilute solutions. It seems therefore as if the current is more easily carried through the solution by a given mass of dissolved substance in a dilute than in a concentrated solution.

The temperature coefficient of resistance of electrolytes is negative, or the conductivity *increases* with the temperature.

$$k_t = k_0 (1 + \alpha t) \qquad (348)$$

The temperature coefficient of conductivity is usually large. For sulphuric acid it is 0.016, for copper sulphate 0.0226 and for zinc sulphate 0.025 per degree.

On account of polarization in the cell electrolytic resistance cannot be measured by the usual Wheatstone-bridge method, as given (Art. 272). The method must be modified by substituting a rapidly alternating current for the constant current, and a telephone receiver for a galvanometer.[1]

288. Practical Applications of Electrolysis. Electrolytic methods are employed on a large scale for the refinement of copper, aluminium and other metals. A large slab of the impure metal is used as the anode, and the pure metal is deposited at the cathode. Nearly all the impurities are separated at the anode, and either remain in solution or collect as a dirty slime at the bottom of the tank.

Electrotyping is a process of taking exact copies of coins, engravings, etc., by depositing some metal, such as copper or zinc, upon an impression of the original object, made in wax, gutta percha or some similar substance. The impression is carefully coated with a good conductor, as graphite or powdered bronze, and used as the cathode of an electrolytic cell.

Electroplating is the process of covering baser metals with precious metals by means of electrolysis. For gold and silver plating the cyanide salts of gold and silver are used as electrolytes. For nickel plating a double sulphate of nickel and ammonium is used. The anode must in all cases be of the metal which is to be deposited at the cathode.

[1] *For methods for the measurement of electrolytic resistance, see Manual, Exercises 61 and 69.*

CHAPTER XXXVI

ELECTRIC CELLS

289. Polarization of a Cell. It was shown (Art. 286) that polarization occurs when the products of electrolysis are chemically different from the electrodes upon which they collect. When a simple voltaic cell furnishes a current, hydrogen appears upon the copper plate (Art. 253), and the resulting polarization interferes with the efficiency of the cell. The principal facts concerning polarization in a primary cell may be illustrated by means of a cell having zinc and *clean, pure* mercury for electrodes and an aqueous solution of common salt as the electrolyte. The zinc plate hangs in the solution, while an insulated wire leads down to the layer of pure mercury which forms the positive plate at the bottom of the cup.

On closing the circuit of this cell through a telegraph sounder of low resistance, the signals are given sharply for a few seconds, then faintly, and finally cease entirely. Owing to the film of hydrogen which is collected upon the mercury surface, the counter E.M.F. of polarization is almost equal to the direct E.M.F. of the cell, and the current is reduced almost to zero. The cell is *polarized.* By dropping into the cell a piece of mercuric chloride as large as the head of a pin, the cell is instantly restored to action.

The mercuric chloride reacts upon the free hydrogen, forming hydrochloric acid (HCl) and free mercury. This disposes of the polarizing film of hydrogen, and the signals are given clearly so long as the mercuric chloride lasts. Any substance which unites chemically with the hydrogen film upon the positive plate, and thus reduces polarization, is termed a *depolarizer.* In the construction of cells, care must be taken either to have

no polarization or to supply the cell with a depolarizer. In the following paragraphs only such cells have been described as are found in common use.

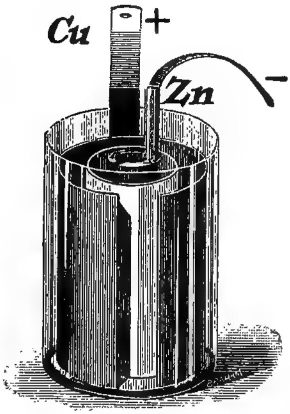


FIG. 149.

290. The Daniell Cell. The Daniell cell consists of a zinc electrode immersed in a dilute solution of zinc sulphate and a copper electrode immersed in a saturated solution of copper sulphate. The two liquids are kept apart by means of a porous cup (Fig. 149). To insure saturation of the copper sulphate solution crystals of copper sulphate are added in excess. This cell has an E.M.F. of 1.09 volts. During the action of the cell zinc goes into solution, and copper is deposited upon the copper plate, since the current *in the cell* flows from the zinc to the copper electrode. There is no polarization, and the E.M.F. of the cell remains constant, while a current is drawn from it.

The gravity cell is a special type of the Daniell cell. In this the porous cup is omitted, the solutions being kept separate by placing the denser copper sulphate solution at the bottom of the cell (Fig. 150). The gravity cell has a much smaller internal resistance than the ordinary Daniell cell, and can thus furnish greater currents.

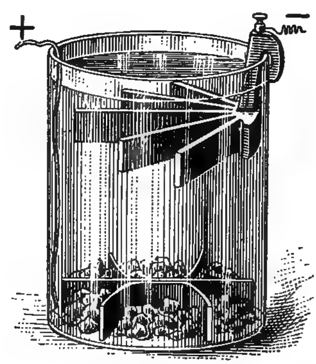


FIG. 150.

The Daniell cell is a *closed circuit cell*, since it must be kept upon closed circuit when not in use. If it be left on open circuit, the copper sulphate diffuses toward the zinc and deposits upon it a muddy mass of copper and copper oxide, interfering with the action of the cell.

*291. **The Bichromate Cell.** The bichromate cell consists of zinc and carbon electrodes immersed in a mixture of a solution of potassium or sodium bichromate and concentrated sulphuric acid. The E. M. F. of the cell is 2.1 volts when fresh, but soon falls to about 1.75 volts. The bichromate acts as an efficient, though imperfect, depolarizer. Since the zinc is constantly dissolved by the sulphuric acid, it must be taken out of the solution when the cell is not in use.

292. **The Leclanché Cell.** In this cell the electrodes consist of carbon and zinc, and the electrolyte is a solution of ammonium chloride. The solid depolarizer is a mixture of manganese dioxide and crushed carbon which is usually packed around the carbon in a porous cup (Fig. 151) or in a canvas bag. The dioxide is a very poor depolarizer, and for this reason the E. M. F. of a Leclanché cell decreases rapidly when on a closed circuit of small resistance, but it soon recovers when the circuit is opened. The E. M. F. of the cell is about 1.5 volts.

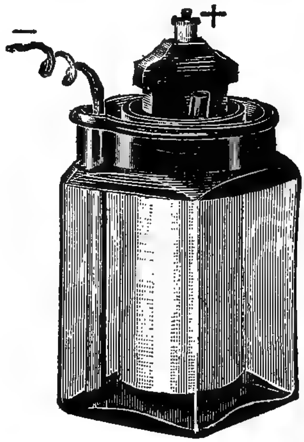


FIG. 151.

The Leclanché cell is an *open circuit cell*, or it can be kept indefinitely upon open circuit without deterioration, and is therefore admirably suited for such intermittent service as that required for door bells, annunciators, electric signals and the like.

The *dry cell* is a modified form of the Leclanché cell. It is not really dry, since the substance between the electrodes is a moist paste, consisting of ammonium chloride, zinc chloride, zinc oxide and plaster of Paris. These cells are very convenient for general use, although their internal resistance soon becomes relatively high.

293. **The Storage Cell.** If two lead plates be placed in a solution of sulphuric acid and a current be sent through the

cell, it will be found that the polarization produced is much more persistent than in the example studied (Art. 286). If the current be continued for a sufficiently long time, the anode becomes covered with a brown coating of peroxide of lead (PbO_2). On disconnecting from the original source of current and closing the electrolytic cell through a resistance, it is found capable of furnishing a current for some time. *The current now leaves the cell by the same electrode through which it entered.* The cell has become a generator of electricity with a voltage of a little more than 2 volts. Any cell in which the energy of an electric current can be changed to such a form that it may, in its turn, be used for the production of electrical energy, is called a *storage cell, an accumulator or a secondary cell.*

The best-known cell of this kind is the lead storage cell, in which the positive electrode is formed of lead peroxide, and the negative of pure lead. The electrolyte is dilute sulphuric acid of density 1.26 grams per cc. Many different types of these cells are on the market, differing mainly in the mode of preparation of the plates and in the manner in which the active material is held in specially designed frames or grids.

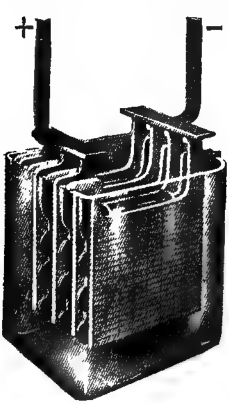


Fig. 152.

Owing to its extremely small internal resistance, this cell is especially useful for furnishing large currents. In the larger forms there are in each cell a number of negative plates, all connected together, and between them the positive plates, also joined together (Fig. 152). The greater the number of plates, the greater is the current carrying capacity of the cell.

While the cell furnishes a current, the peroxide coating is slowly reduced to lead sulphate, while the lead plate becomes oxidized to the same form. When, finally, the two electrodes are electrochemically alike, the E. M. F. drops off rapidly and the cell is discharged. In order to restore it to its original condition, it is only necessary to send a current through it in a direction opposite to that

obtained when the cell is in use. This process is called charging. The whole process may be represented by the chemical equation

$$PbO_2 + 2\,H_2SO_4 + Pb \rightleftarrows 2\,PbSO_4 + 2\,H_2O$$

The storage cell is of great commercial importance, and large batteries are installed in many electric power plants. During certain hours of the day the demand for electric power is far below the average, and during this time the cells may be charged from the generators. After being charged, they are used to supply current during the hours of maximum demand.

Edison has recently placed upon the market a storage cell of a different type. The active materials consist of nickel oxide, NiO_2, held in a steel grid, forming the positive electrode, and finely divided iron, placed in pockets of the negative electrode. The electrolyte is an aqueous solution of caustic potash of specific gravity 1.20. The containing jar is made of nickel plated sheet steel. The E.M.F. of this cell is 1.2 volts. The electrical energy furnished by these cells upon discharge is about twice as large as that furnished by a lead cell of the same weight. The electrochemical process in these cells is not well understood. It is probably about as follows:

$$Fe + 2\,NiO_2 + H_2O + KOH \rightleftarrows FeO + Ni_2O_3 + H_2O + KOH$$

*** 294. Energy Relations.** While a cell furnishes a current, certain chemical reactions are going on in the cell, and the energy set free by these chemical reactions is, in part at least, transformed into electrical energy (Art. 253). Assuming that all this energy is obtained as electrical energy, we may calculate at once the E.M.F. of the cell. Consider, for example, the Daniell cell, in which the chemical reaction is extremely simple. Zinc goes into solution and copper separates out. The resultant reaction, therefore, consists in replacing one chemical equivalent of copper in a copper sulphate solution by one of zinc, thus forming a zinc sulphate solution. The energy rendered available by this exchange is simply the amount of energy by which the heat of formation of zinc sulphate exceeds that of copper sulphate. This excess amounts to 25,000 calories.

If this process take place as a purely electrochemical reaction, these 25,000 calories are all rendered available as electrical energy. The corresponding electrical energy is EQ, where E is the E. M. F. of the cell, and Q is 96,530 coulombs (Art. 284). Therefore

$$96{,}530\,E = 25{,}000 \text{ calories} = 4.2 \times 25{,}000 \text{ joules} \qquad (349)$$

$$E = \frac{105000}{96530} = 1.09 \text{ volts} \qquad (350)$$

In many cells the above assumption does not hold, since temperature changes occur at the electrodes during the passage of the current. For a more complete discussion of these energy relations, the student is referred to more advanced texts.

295. Fall of Potential in a Circuit containing a Cell.[1] It is now generally assumed that the E. M. F. of a cell has its seat at the surface between the electrodes and the electrolyte, and is therefore the sum of two single potential differences. These quantities cannot be measured independently.

The Daniell cell may be taken as an example to illustrate the fall of potential around the entire circuit. The zinc, being the negative electrode, has the lowest potential. Passing from the zinc plate to the electrolyte, there is a sudden rise of potential at the surface of the zinc, and a second rise on passing from the electrolyte to the copper. If the cell be closed through an external resistance, there is a fall of potential equal to Ir in the cell, where r denotes the internal resistance of the cell. In the external resistance R there is a fall of potential equal to IR; and according to Kirchhoff's second law, their sum must be equal to E, the E. M. F. of the cell.

Plotting the potentials along the circuit as a function of the resistance, we obtain a figure similar to the one shown (Fig. 153), where A represents the potential of the zinc, AB the rise of potential at the surface of the zinc, $B'C$ the drop in the cell, CD the rise at the surface of the copper and FD the fall of potential through the external circuit, back to A', the potential of the zinc. The figure should be thought of as

[1] *For experimental verification see Manual, Exercise 66.*

a complete circuit, so that the point A' coincides with A (Fig. 154).

AG (Fig. 153) is evidently the E. M. F. of the cell. The actual difference of potential FD, between the electrodes is called the *terminal potential difference*, which, on a closed circuit, is always smaller than the E. M. F. of the cell. The two are equal only when the cell is on open circuit, or when I is zero. The line $ABB'D'$ then represents the change of potential in the cell.

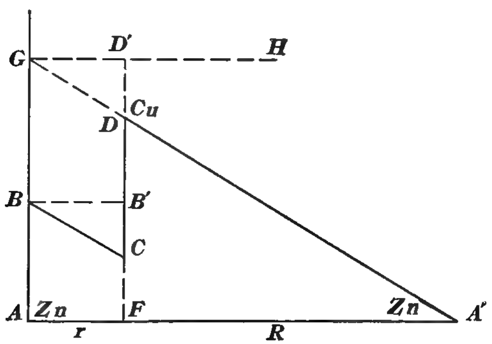


Fig. 153.

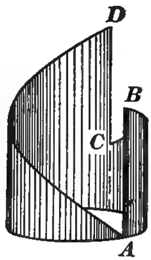


Fig. 154.

296. Cells in Series. When it is desired to obtain an electromotive force larger than that furnished by a single cell, several cells may be placed in series, so that the negative pole of one cell is joined by a conductor to the positive pole of the next, etc. (Fig. 155). In this case the electromotive forces as well as the internal resistances are added.

Fig. 155.

Let such a battery consist of n cells in series, each cell having an E. M. F. of E volts and an internal resistance of r ohms; then by Kirchhoff's second law we have

$$I = \frac{nE}{nr + R} \tag{351}$$

If R be so large that nr may be neglected, we have

$$I = \frac{nE}{R} \tag{352}$$

or the current is n times as large as that which would be furnished by a single cell through the same external resistance.

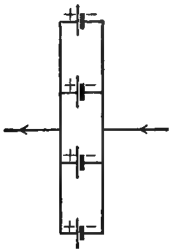


FIG. 156.

297. Cells in Parallel. When a number of cells are so joined up that all the negative poles are connected to one terminal of the circuit and all the positive poles to the other, the cells are said to be connected in parallel (Fig. 156). In this case the E. M. F. of the battery does not rise above that of a single cell, while the internal resistance has been reduced to $\frac{1}{n}$th of that of a single cell, in accordance with the law of parallel circuits (Art. 277).

If n cells, each of E. M. F. E and internal resistance r, be connected in parallel, and the external resistance be R ohms, then on closing the circuit the current is

$$I = \frac{E}{\frac{r}{n} + R} = \frac{nE}{r + Rn} \tag{353}$$

If nR be negligible in comparison with r, this arrangement will give a current n times as large as that from a single cell through the same external resistance. This arrangement is therefore to be preferred in all cases where the resistance of the battery forms the major part of the resistance to be overcome.

298. Standard Cells. The cells described in this article are not intended to furnish current. Their great importance lies in the fact that they have a definite and constant E. M. F. on open circuit, and that they may be used as concrete standards for the measurement of differences of potential.

(*a*) The *Clark standard cell.* This cell (Fig. 157) consists of an H-shaped, hermetically sealed glass tube, containing pure mercury as the positive electrode in one leg, and a zinc amalgam of 10 to 15 per cent of zinc as negative electrode in the other. Platinum wires sealed through the glass connect the electrodes with the circuit. Above the mercury is placed a paste of mercurous sulphate, while the electrolyte is a sat-

urated solution of zinc sulphate. In order to insure saturation at all temperatures, an excess of zinc sulphate crystals is added.

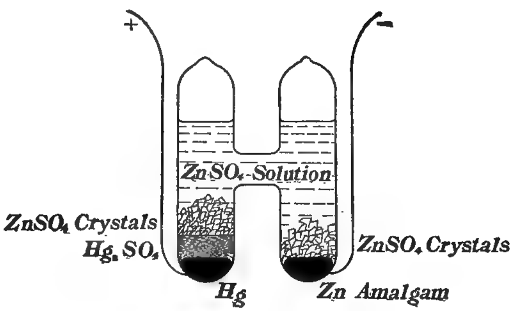


FIG. 157.

The E. M. F. of the Clark standard cell is 1.433 volts at 15° C, and at any other temperature $t°$, it is given by the equation

$$E_t = 1.433 - 0.00119\,(t - 15) \tag{354}$$

(*b*) The *cadmium or Weston standard cell* is of the same type as the Clark cell, except that cadmium and cadmium sulphate are used instead of zinc and zinc sulphate. The variation of E. M. F. with temperature is very small, and at present this cell is considered the best standard cell available. Its E. M. F. at $t°$ is

$$E_t = 1.0183 - 0.00004\,(t - 20) \text{ volts} \tag{355}$$

299. Definition of the Volt. "*The international volt is the electric pressure which, when steadily applied to a conductor whose resistance is one international ohm, will produce a current of one international ampere*" (Art. 264). The volt is thus defined in terms of the ohm and the ampere in accordance with Ohm's law. At the same time, however, it is of advantage to express the volt in terms of some concrete standard. To this end the value of the E. M. F. of the Weston standard cell, in terms of the international volt, as given in equation (355), has been determined by a number of very careful experiments. We may therefore provisionally take the volt as $\frac{10000}{10183}$ of the E. M. F. of the Weston standard cell, at 20° C.

CHAPTER XXXVII

THERMOELECTRICITY

300. The Seebeck Effect. Seebeck[1] discovered in 1822 that a current flows through a circuit formed of two different metals when the two junctions are at different temperatures. Such a pair of metals is called a thermocouple or a thermoelement.

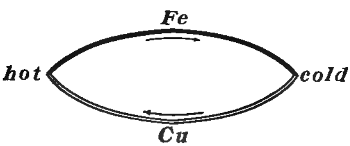


FIG. 158.

Thus, in a couple of iron and copper (Fig. 158) the current at the hot junction flows from the copper to the iron. The electrical energy which appears in this phenomenon is evidently derived from a small quantity of heat which has been absorbed and transformed by the thermocouple.

301. The Peltier Effect. The reverse phenomenon was discovered by Peltier[2] in 1834, namely, that a cooling or heating of the junctions of two dissimilar metals occurs when an electric current passes through them. Thus, at a copper-iron junction heat is *absorbed* when the current passes from the copper to the iron, and heat is *evolved* when the current passes from the iron to the copper. The thermal effect is proportional to the quantity of electricity flowing through the junction.

This phenomenon is explained by the existence of electromotive forces at the surfaces of separation of the two metals. If we call the sum of these two E.M.F.'s E, the energy EQ, necessary to force the electric quantity Q through the two junctions, must be equal to the heat absorbed at one junction minus the heat liberated at the other. Where the heat is absorbed, work is being done upon the electric current, or the current is made to flow from a lower to a higher potential.

[1] Seebeck, *Abh. Ak. Wiss. Berlin*, 1822–23, p. 265.
[2] Peltier, *Ann. Chim. et Phys.*, 56, p. 371, 1834.

This junction is an electric generator and may be compared to a cell, the iron and the copper forming the positive and the negative terminal respectively. Heat is transformed into electric energy and the current passes from lower to higher potential, through the heated junction, just as it passes from the negative zinc to the positive copper in the voltaic cell. The electromotive force at the junction is called the *Peltier* E.M.F. Where heat is evolved, the electrical energy decreases, or the current flows from a higher to a lower potential, against the Peltier E.M.F. at this junction.

* **302. The Thomson Effect.**[1] If a uniform metallic rod (Fig. 159) be heated at one point A, heat will be conducted at the same rate from this point to either side, towards B and C. But if at the same time an electric current be sent through the metal, the resulting flow of heat is no longer the same on the two sides of A. According to the direction of the current, the flow of heat is smaller on one side and larger on the other. This is equivalent to saying that energy in the form of heat is absorbed on one side of the heated point and liberated on the other.

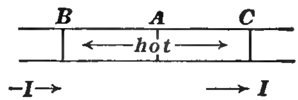


FIG. 159.

This phenomenon is readily explained by assuming that a difference of potential exists between the hotter and the cooler portions of the conductor. To fix our ideas, let us suppose the hot point A to be at the higher potential, and the current to be sent from B to C. In the section BA the current flows from lower to higher potential, and work must be done upon it, or heat is absorbed. In the section AC the current flows from higher to lower potential, and electric energy is transformed into heat. The final result is therefore an apparently larger flow of heat from A towards C than from A towards B, or a flow of heat *with* the current. Among other metals, copper, antimony and silver show this effect as described. On the other hand, the effect is reversed in the case of a large number of metals, such as iron, bismuth, tin and platinum. In these metals the increased flow of heat is in the opposite direction to

[1] Thomson, *Phil. Trans.* 1856, 3, p. 661.

that of the electric current, and the potential gradient must be assumed to be opposed to the temperature gradient.

Both *BA* and *AC* contain electromotive forces which are called *Thomson E.M.F.'s*. In *BA*, for example, the positive terminal is at *A*, the negative at *B*, and its E. M. F., if present alone, would tend to send a current from *B* to *A*. In the portion *AC* the effect is reversed. In a copper-iron element, with its junctions at different temperatures, the Thomson E. M. F., if present alone, would send the current at the hotter junction from the copper to the iron. In a closed circuit, consisting of one metal only, heated at some point in the circuit the two opposed electromotive forces are equal, and no electric current, due to unequal heating of the circuit, can be observed.

303. Thermoelectromotive Force. The Seebeck effect (Art. 300) in a circuit, consisting of two homogeneous metals, may be considered as due to the sum of the two Peltier E. M. F.'s at the *junctions* and the two Thomson E. M. F.'s along the *conductors*. This total E. M. F., arising from a temperature difference at the junctions of two dissimilar substances, is called a *thermoelectromotive force*.[1]

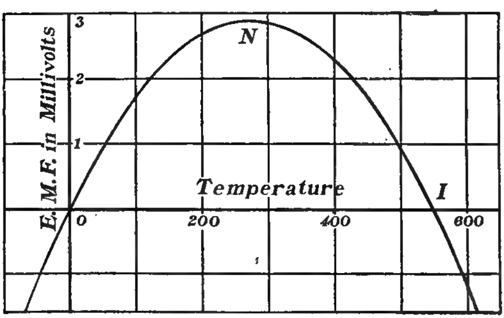


FIG. 160.

The relation between temperature and thermoelectromotive force in the case of a copper-iron element, where one junction is kept at 0° C, and the temperature of the other is varied, is shown (Fig. 160). The E. M. F. increases with the temperature, but at a decreasing rate, until it reaches a maximum at 275°, the so-called *neutral temperature*. Beyond this point the E. M. F. decreases and reaches a zero value at 550°, which is called the *temperature of*

[1] *For the determination of thermoelectromotive force, see Manual, Exercise 65.*

inversion. If the hotter junction be at a temperature above 550°, the thermoelectromotive force is reversed, or the current flows from the iron to the copper through the hotter junction.

Let a thermoelement show an E. M. F. E_1, when one of its junctions is at 0° C and the other is at t_1°; also let E_2 be the E. M. F. of the same element when the first junction is at 0° and the second at t_2°; then if the temperature of the first junction be raised to t_1°, while that of the second is kept at t_2°, the resulting E. M. F., E, is given by the equation

$$E = E_2 - E_1 \tag{356}$$

or thermoelectromotive force is an additive quantity.

Though we are here studying only thermoelectromotive forces in a circuit, consisting of metallic conductors, it should be mentioned that similar effects are observed at the surfaces between metals and electrolytes.

***304. Thermoelectric Power.** The variation of the thermoelectromotive force per degree centigrade is called the thermoelectric power, and may be represented by

$$P = \frac{E'' - E'}{t'' - t'} = \frac{dE}{dt} \tag{357}$$

where t'' and t' are very close together. If the thermoelectric power for any couple be plotted as a function of temperature, a straight line is obtained, passing through the temperature axis at the neutral temperature (Fig. 161).

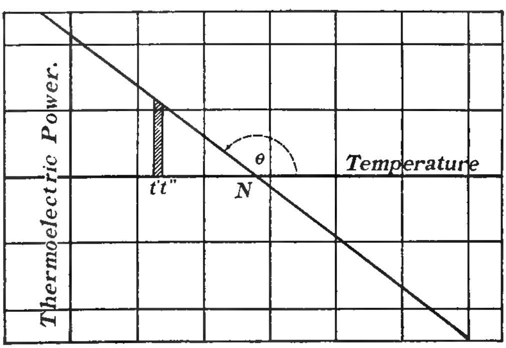


FIG. 161.

Equation (357), if written in the form

$$E'' - E' = P\,(t'' - t') \tag{358}$$

shows that the E. M. F. between the temperatures t'' and t' is given by the narrow strip whose height is the thermoelectric power and whose width is $t'' - t'$. From equation (358) it is seen that the E. M. F. for any temperature difference whatsoever is measured by the area included between the straight line, the two ordinates, and the axis of temperature. An area below this axis must be considered as negative. Thus, if one junction be kept at any temperature below the neutral temperature, and the temperature of the other junction be increased, the thermoelectromotive force of the couple increases until the second junction reaches the neutral temperature, then decreases and becomes zero when the hotter junction is at a temperature as far above the neutral temperature as the colder junction is below it. Beyond that temperature a reversal of the E. M. F. takes place.

The thermoelectric power between practically all metals and alloys, when plotted with temperature, gives similar straight lines, but inclined under a different angle and with different neutral temperatures.

***305. Thermoelectric Series.** At a given temperature the thermoelectric power of any metal A with respect to another metal C is equal to the sum of the thermoelectric powers of A with respect to a third metal B, plus that of B with respect to C. We may therefore arrange all metals in a thermoelectric series, taking one specific metal as a reference standard, and giving the thermoelectric powers of the other metals with respect to this. Such a table is given below for the thermoelectric powers at 20° C, all referred to lead as a standard. In order to obtain the thermoelectric power between any two metals in the series, the values given in the second column must be subtracted one from the other. When the junction of any pair is moderately heated, the current flows through the junction from the metal standing first in the series to the one standing below.

TABLE XVIII

THERMOELECTRIC POWERS IN MICROVOLTS PER DEGREE, AT AN AVERAGE TEMPERATURE OF 20° C

SUBSTANCE	MICROVOLTS	SUBSTANCE	MICROVOLTS
Bismuth	−89	Zinc	+3.7
Cobalt	−22	Copper	+3.8
German silver	−12	Iron	+17.5
Lead	0	Antimony	+24
Silver	+3	Selenium	+807

A very convenient form of thermoelement for low temperature measurements is a copper-constantan couple, whose thermoelectric power is about 170 microvolts per degree C.

306. The Thermopile. Table XVIII shows that the thermoelectromotive force between metals is very small in comparison with the E. M. F. of an electric cell. It is, however, possible to obtain an E. M. F. comparable with that of a cell, if a number of such couples be connected in series (Fig. 162), and the alternate junctions *A*, *B*, *C*, etc., be heated, while the remaining junctions *A′*, *B′*, etc., are kept at a lower temperature.

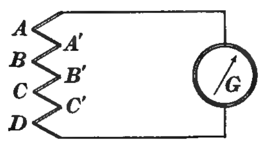


FIG. 162.

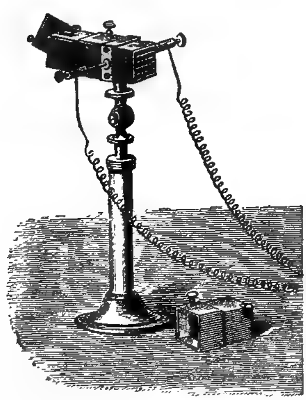


FIG. 163.

Frequently these couples are arranged to form a block of rectangular cross section, held in a suitable case (Fig. 163). Such a set of couples is called a thermopile.

CHAPTER XXXVIII

APPLICATION OF THE HEATING EFFECT OF CURRENTS

307. Electric Heating. The production of heat by an electric current has assumed considerable practical importance, and on account of its great convenience this method of heating is being used more and more extensively. Thus, we have electric cooking, electric ironing, electric soldering, electric welding, etc.

In an electric furnace the current is employed to produce high temperatures. It flows in coils of wire or flat strips embedded in the mass of the furnace or surrounding it. The conductors must be of non-oxidizable material of high melting point. By varying the current, the temperature of the furnace may be adjusted to any desired value and be kept constant without appreciable change.

Since the heating effect of a given current is proportional to the resistance, the temperature will be highest in that part of the circuit which has relatively the greatest resistance. Thus, a thin platinum wire inserted in a circuit may be heated to white heat by a current which hardly affects the temperature of the rest of the circuit. Such local heating of a small piece of the circuit is used in cautery and in electric blasting.

If the piece of inserted wire have high resistivity and low melting point, and the current rise beyond a certain value, the wire will melt and break the circuit, protecting the rest of the circuit from an excessive current. Such wires are called safety fuses, and are used in different sizes according to the maximum current for which the circuit is intended.

308. The Incandescent Lamp. Any conductor may be heated by an electric current to incandescence, and in this condition

may serve as a source of light whose intensity may easily be controlled by adjusting the current strength.

In an ordinary incandescent lamp a thin carbon filament, inclosed in an exhausted glass globe, is heated to bright yellow heat. A 16 candle power lamp designed to be used on a 100-volt circuit has a resistance, when hot, of about 160 ohms, and takes, therefore, a current of about 0.6 ampere. It is called a 60 watt lamp, and its efficiency is given as 3.75 watts per candle.

Lately metallic filaments are being used instead of carbon. The best lamps of this type are the tungsten lamps, in which the heated conductor consists of a very fine wire of an alloy of tungsten and osmium. Both these metals have very high melting points, and may be heated to a higher temperature than carbon, thus giving a whiter light. The efficiency of these lamps is high, since they require not more than 1.5 watts per candle. One disadvantage of the metallic filaments is that they are very brittle, and are therefore quite easily broken by mechanical shocks.

All incandescent lamps are designed **for a** definite voltage. A change of only one per cent from this voltage causes a change of from five to six per cent in the candle power. However, too high a voltage decreases materially the life of the lamp.

309. The Arc Lamp. If two rods of carbon, connected to a source of an electric current, of an E. M. F. higher than 50 volts, be brought into contact, and then separated by a short distance, a brilliant white light appears between them. The electric current is not interrupted by the separation of the carbons, since the space between them has been made conducting (Art. 420). The luminous path of the current has a curved form (Fig. 164), and is therefore called an electric arc. Both carbons, especially the positive, grow very hot, and are the

FIG. 164.

main sources of the light. On the positive carbon a hollow or crater is formed, while the negative carbon becomes conical in shape. The efficiency of the simple arc light is about one watt per candle.

Since the material of the electrode slowly wastes away, due to oxidation in the air, the arc lamp usually contains an automatic regulator, which, after the arc is started, maintains the carbons at a fixed distance apart. If the arc be interrupted, this mechanism allows the carbons to come into contact, and then separates them as soon as the current starts between them.

The positive carbon wastes away twice as fast as the negative, and is frequently made of larger diameter. By inclosing the arc in a small glass globe, supplied with a cover, the rate of consumption is materially decreased, but the luminous efficiency is also greatly reduced. Such lamps are called inclosed arc lamps. The temperature of the electric arc is about 3600° C, the highest temperature yet produced by artificial means. All metals may easily be volatilized in the arc.

The *flaming arc*, now frequently employed for illumination, gives a more brilliant light than the ordinary arc. Besides, its efficiency is as high as 0.3 watt per candle. The carbons in this type of lamp contain a core, consisting of carbon, mixed with lime, magnesia and other oxides which are very efficient sources of light at high temperatures. In these lamps the arc itself is the chief source of light rather than the heated carbon tips.

*** 310. The Nernst Lamp.** The light-producing substance in the Nernst lamp is a narrow cylinder of rare earths similar in composition to that of the Welsbach mantle. This part of the lamp is called the glower. Since these substances do not conduct electricity at ordinary temperatures, they must first be heated. This is done by sending the current through fine platinum wires stretched parallel to the glower and embedded in porcelain. As soon as the temperature of the glower is raised sufficiently by the radiation from the incandescent heater, the current begins to pass through the glower itself, makes it white-hot, and the heater is cut out of the circuit by a separate electromag-

netic mechanism. The action of the Nernst lamp is of an electrolytic nature, and it can therefore be used to advantage only on an alternating current circuit. The efficiency of this lamp is only about 2 watts per candle, and its construction is so complicated that it is rapidly being displaced by the tungsten lamp.

311. The Cooper-Hewitt Lamp. In this lamp a column of mercury vapor in an exhausted glass tube, two or three feet long, forms the conductor of the electric current. The lamp is held in a slightly inclined position. A plate of iron at the upper end of the tube forms the positive electrode, while the negative electrode at the bottom is a small amount of mercury. In order to start the lamp, the tube is slightly tilted so that the mercury comes in contact with the iron plate and closes the circuit. When the lamp is returned to its original position, an arc appears at the point where the circuit breaks, and the whole mass of mercury vapor in the tube becomes incandescent, emitting a greenish blue light. The efficiency of this vapor lamp is very high, as it requires but 0.3 watt per candle. But, since the light is deficient in red and yellow rays, it gives all colored bodies an unnatural appearance. Its rays are, however, very effective for photography, and these lamps are much used for that purpose.

CHAPTER XXXIX

ELECTRICAL CONDENSERS

312. Action of the Condenser. An electrical condenser (Fig. 165) consists of a large number of sheets of tin foil, separated by thin sheets of insulating material, such as paraffined paper or mica. The tin foil is arranged in two sets, so that each sheet lies between two of the other set. Each set is connected to a binding post, the two posts A and B thus forming the terminals of the condenser.

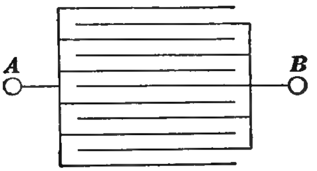


FIG. 165.

The action of a condenser is easily shown by the following experiment, in which a condenser C (Fig. 166), a cell E and a galvanometer G, having a movable system of large moment of inertia, are connected as shown. The condenser is first connected in series with the cell by closing the key k at A; an instant later the circuit is closed through the galvanometer by throwing the key over to B. A deflection of the galvanometer is obtained which reaches a maximum and then returns to zero. This throw of the galvanometer is due to a transient current, that is, the passage through the instrument of a quantity of electricity which was stored in the condenser while it was connected to the cell. This was released when the condenser was short circuited through the galvanometer circuit which contained no E. M. F. The condenser is said to have been charged by the cell and discharged through the galvanometer.

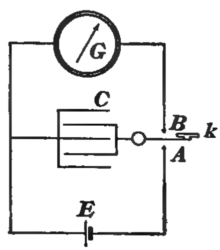


FIG. 166.

313. Capacity of a Condenser. If the experiment of the previous article be repeated with different electromotive forces, it will be found that the quantity of electricity stored in a given condenser is always proportional to the E. M. F. to which the condenser is connected, or, in mathematical terms,

$$Q = CE \tag{359}$$

where C is a characteristic constant of the condenser, and is called the *capacity of the condenser.*

It may be defined as the quantity stored in the condenser per unit difference of potential of the charging source. It is numerically equal to the quantity which produces unit difference of potential at the terminals of the condenser.[1]

314. Unit of Capacity. *Unit capacity is that capacity which is charged to unit difference of potential by unit quantity of electricity.* In practical units, unit capacity is that capacity which is charged to a difference of potential of one *volt* by one *coulomb.* This unit is called the *farad*, after the English physicist, Faraday (1791–1867). Since the farad is far greater than the capacity of ordinary condensers, the unit in common use is the *microfarad*, or the one millionth of a farad. Thus the quantity of electricity stored in a large commercial condenser of three microfarads capacity, by a difference of potential of one hundred volts, is only 0.0003 coulomb. The capacity of ordinary metallic circuits is quite small, but it should not be assumed that only condensers such as those described (Art. 312) have capacity. Every conductor has a capacity, especially when other conductors are in the neighborhood. The capacity of underground circuits is often quite large, while the capacity of submarine cables frequently exceeds one microfarad per kilometer.

315. Mechanical Analogue. The presence of an insulator or dielectric in a condenser prevents the passage of a continuous current through the circuit. But during the charging or discharging of a condenser there is a transient flow of electricity through the conductor. As in Art. 266, we may compare an

[1] *For experimental determination of capacity, see Manual, Exercises 72 and 73.*

electromotive force to a difference of pressure produced by a pump in a system of pipes. If there be no obstruction, there will be a continuous circulation of fluid through the pipes. But if a number of elastic membranes *a* to *f* (Fig. 167) be stretched across the pipe in some part of the circuit, the pump *P* will merely produce a small displacement of the fluid in the system until the back pressure due to the elastic reaction of the membranes becomes equal to the pressure exerted by the pump. If now the pump be removed, the strained membranes will cause a counterflow through the pipes.

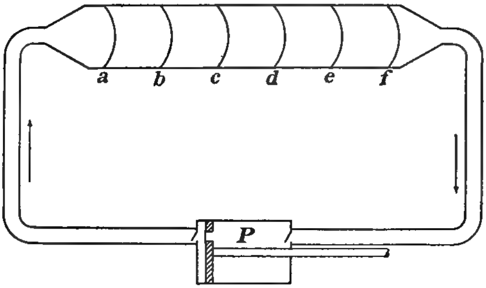


FIG. 167.

The elastic membranes represent the dielectric of the condenser, and we may consider the electricity stored in the condenser as equivalent to an elastic displacement in the insulating material. From this point of view the electric quantity charging the condenser does not stop at the surfaces of separation between the conductors and the dielectric of the condenser, but equal quantities are displaced through the whole circuit. Maxwell, who held this view, speaks, therefore, of *displacement currents* in the dielectric.

316. The Dielectric Constant. If two conducting plates of equal area A be separated by a thin sheet of a dielectric of thickness d, the capacity of such a condenser is found to be very nearly proportional to the *area* of one of the plates, and inversely proportional to the *distance* between them. It also depends greatly upon the nature of the dielectric. The capacity of such a condenser may be computed from the formula

$$C = \frac{885}{10^{10}} \cdot c \cdot \frac{A}{d} \text{ microfarad} \qquad (360)$$

where A and d are expressed in C. G. S. units. The constant c, which depends upon the dielectric, is called the *dielectric constant*, or inductivity of the substance of the dielectric. The dielectric constant of a vacuum is arbitrarily chosen as unity, but it is also practically unity for air.

We may therefore measure the dielectric constant of any substance by comparing the capacity C_a of a condenser, in which air forms the dielectric, with its capacity C_x when the substance in question forms the dielectric between the condenser plates. For, in this case,

$$\frac{c}{1} = \frac{C_x}{C_a} \tag{361}$$

or, the dielectric constant is measured by the ratio of the two capacities.

The dielectric constant of paraffine is 2, of sulphur and ordinary glass 3, of mica 6 and of flint glass 8.

The *dielectric strength* is the maximum difference of potential per centimeter thickness which an insulating material can support without rupture. The following are approximate values for dielectric strength: paraffine oil, 87,000; solid paraffine, 130,000; beeswaxed paper, 540,000 volts per centimeter.

317. Electric Absorption and Leakage. If a commercial condenser be discharged, it will be found that the amount of electricity depends somewhat upon the time of charging. The condenser will "absorb" electricity which seems to pass into the dielectric. After such an absorbing condenser is discharged and disconnected from the discharging circuit, the absorbed charge slowly reappears at the conducting plates, and thus a number of consecutive discharges of constantly decreasing quantity may be obtained.

The actual capacity of commercial condensers is therefore not a very definite quantity, but depends upon the rate of charge and discharge. Mica shows but little absorption, and is therefore used in the construction of standard condensers. Many condensers also show a *leakage*, or passage of electricity from one terminal to the other, thus indicating poor insulation.

The resistance of a condenser should be at least several million ohms, even in the poorest condensers.

318. Condensers in Parallel and in Series. If a number of condensers of capacities C_1, C_2, C_3, etc., be placed in parallel (Fig. 168), the whole system will have a capacity C equal to the sum of the separate capacities. For, since they are all charged to the same difference of potential E, the total quantity of electricity stored in the system is

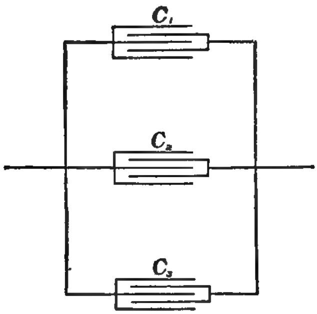


FIG. 168.

$$Q = EC = Q_1 + Q_2 + \cdots$$
$$= EC_1 + EC_2 + \cdots \qquad (362)$$

or $$C = C_1 + C_2 + C_3 + \cdots \qquad (363)$$

If these condensers be placed in series (Fig. 169), the same quantity of electricity is stored in each, since the same displacement current passes every cross section of the circuit, while the difference of potential between the terminals of the different condensers will, in general, be different. We obtain, therefore, in this case the equations

FIG. 169.

$$EC = Q = E_1C_1 = E_2C_2 = E_3C_3 = \cdots \qquad (364)$$

also $$E = E_1 + E_2 + E_3 + \cdots \qquad (365)$$

whence, substituting E_1, E_2, etc., from (364) we have

$$\frac{Q}{C} = Q\left(\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \cdots\right) \qquad (366)$$

or, finally, $$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \cdots \qquad (367)$$

Problems

1. In what time will a constant current of one ampere decompose one gram of water? *Ans.* 2 hr, 58 min, 45 sec.

2. A current passes through three electrolytic cells in series; the first contains a solution of silver nitrate, the second a solution of copper sulphate and the third dilute sulphuric acid. It is found that 2.7 g of silver are deposited. Calculate the masses of copper, hydrogen and oxygen liberated.
Ans. Copper, 0.7957 g; hydrogen, 0.0250 g; oxygen, 0.2002 g.

3. Two electrolytic cells, each containing copper sulphate and having the same resistance, are placed first in series and then in parallel for the same length of time. Compare the total quantities of salt decomposed in the two cases, if the E. M. F. in both cases be the same. *Ans.* As 1 to 2.

4. The E. M. F. of a battery is 10 volts. When producing a current of 5 amperes, the terminal potential difference is 8 volts. Find the internal resistance of the battery. *Ans.* 0.4 ohm.

5. A battery having an E. M. F. of 10 volts sends a current through two electrolytic cells, arranged in series, which offer a counter E. M. F. due to polarization equal to 1.5 volts in each cell. Compute the current through the circuit, if the resistance of the battery be 0.5 ohm, that of each cell 2.25 ohms and that of the rest of the circuit 9 ohms. *Ans.* 0.5 ampere.

6. A cell gives a current of 1 ampere when its terminals are joined by a wire of no appreciable resistance, and 0.4 ampere when joined by a wire of 2 ohms resistance. Find the E. M. F. and the internal resistance of the cell. *Ans.* (*a*) 1.333 volts.
(*b*) 1.333 ohms.

7. Twelve cells, each of an E. M. F. of 1 volt and an internal resistance of 2 ohms, are connected to an external resistance of 10 ohms. Compute the current through the circuit (*a*) if the cells be joined in series, (*b*) in parallel, (*c*) in 3 parallel rows, each containing 4 cells in series.
Ans. (*a*) 0.353 ampere; (*b*) 0.098 ampere; (*c*) 0.316 ampere.

8. A storage cell having a resistance of 0.02 ohm develops a counter E. M. F. of 2.5 volts while it is being charged by a current of 5 amperes for two hours. During the discharge it gives a current of 9 amperes for one hour, while its average E. M. F. during this time is 2 volts. What is the electrical efficiency of the cell, considering only the available energy at the terminals? *Ans.* 63 per cent.

9. Compute the E. M. F. of a thermopile, consisting of 150 couples of bismuth and antimony, when one side of the pile is heated to 500° C, while the other is kept at 20° C, assuming the thermoelectric power to be constant in this interval. *Ans.* 8.136 volts.

10. What is the resistance of a 60-watt tungsten lamp, to be used on a 100-volt circuit? What is its candle power? What are the corresponding values for a 60-watt carbon filament?

Ans. Tungsten, 166.7 ohms, 40 c. p.; carbon, 166.7 ohms, 16 c. p.

11. Compare the cost of ordinary incandescent lamps with that of tungsten lamps, when burned for 80 hours on a 100-volt circuit, each set giving a total illumination of 200 candle power, assuming the price for electrical energy to be 12.5 cents per kilowatt hour.

Ans. Tungsten, $3.00; carbon, $7.50.

12. Calculate the capacity of a condenser consisting of 200 sheets of tin foil, each of area 200 cm^2 and separated by mica sheets 0.05 mm thick.

Ans. 4.227 microfarads.

13. Three condensers have capacities of 1, 0.4 and 0.1 microfarads respectively. What will be the capacity (*a*) if they be connected in parallel, (*b*) in series? *Ans.* (*a*) 1.5 microfarads; (*b*) 0.074 microfarad.

14. Compute the quantity of electricity stored in a system of two condensers of 3 microfarads each, when connected to an E. M. F. of 6 volts, (*a*) in parallel, (*b*) in series. *Ans.* (*a*) 36×10^{-6} coulomb.
(*b*) 9×10^{-6} coulomb.

CHAPTER XL

ELECTROMAGNETICS

319. Magnetic Effect of a Solenoid. A solenoid is a helical coil of wire of many turns, carrying a current, and usually of cylindrical form. Each turn of the solenoid produces a magnetic field in the same sense with the result that a strong

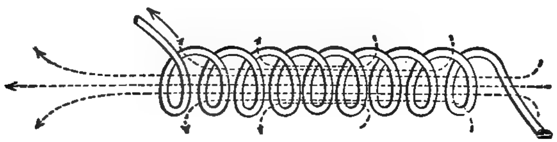


FIG. 170.

magnetic field is produced inside the coil. The direction and sense of the field (Fig. 170) may easily be found by application of the right-hand rule (Art. 256).

The magnetic intensity in the interior of a solenoid whose length is great in comparison with its cross section may, by the use of calculus,[1] be shown to be

$$H = 4\,\pi n I \text{ gauss} \tag{368}$$

where n denotes the number of *turns per unit length* of the solenoid, and I the current strength in C.G.S. units. If I be expressed in amperes,

$$H = \frac{4\,\pi n}{10} I \text{ gauss} \tag{369}$$

This equation also holds for solenoids bent so as to form a closed ring. The magnetic field produced by a ring solenoid is restricted entirely to the closed space inside *the spiral* forming the ring.

[1] See Foster and Porter, *Electricity and Magnetism*, 3d edition, p. 363.

320. Electromagnets. A piece of soft iron placed inside a solenoid (Fig. 171) becomes powerfully magnetized. Such a magnet

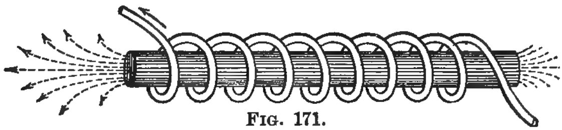


FIG. 171.

is termed an *electromagnet*, and is much more powerful than the permanent magnets which have been studied in previous chapters.

FIG. 172.

The iron, however, loses a large part of its magnetism the instant the current is interrupted. In accordance with the molecular theory (Art. 239), the minute molecular magnets lose their alignment as soon as they are freed from the directing force of the magnetizing field.

In electromagnets of the horseshoe type (Fig. 172), the wire is so carried round the two legs of the magnet as to make the winding continuous if the bar were straightened out. This winding brings opposite poles near to each other, and renders both poles available for lifting or holding by means of the magnet.

Large electromagnets of this type have an enormous lifting power, and are frequently attached to electric cranes in iron foundries. They are able to hold huge masses of iron or steel while these are carried from one part of the building to another.

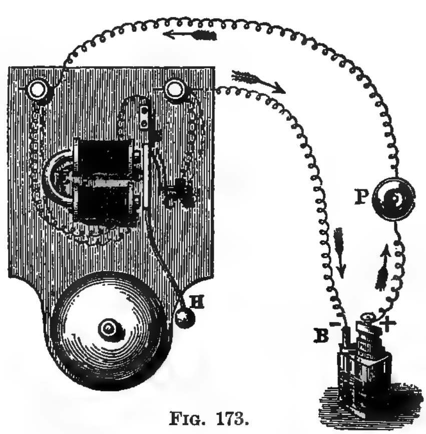


FIG. 173.

A familiar application of the electromagnet is found in the electric bell (Fig. 173). The hammer H, pressed against the point C by the spring s, makes

electric connection with one of the terminals of the cell *B*. The soft iron armature *c*, attached to the hammer, is actuated by the electromagnet, when the electric circuit is completed by the push-button *P*. The armature, when attracted towards the magnet, breaks the circuit at *C* and is drawn back against *C* by the spring *s*. The circuit is thus automatically made and broken, and a rapid vibration of the hammer results. Tuning forks may be kept in continuous vibration by a similar device.

* **321. The Electric Telegraph.** A most important application of the electromagnet is found in the electric telegraph. In its simplest form a telegraphic system consists in some device for the transmission of a set of prearranged signals, denoting letters, words or phrases. In the electric telegraph this transmission is effected by means of a circuit containing an electromagnet, in which the current is made and broken by means of an interrupter or key. To this end a battery is needed and an insulated metallic line connecting the two distant stations. The earth serves as the return circuit. At each end of the line is located a key and a sounder or receiving instrument.

The sounder (Fig. 174) consists simply of a strong electromagnet with a pivoted armature moving between two narrow detents or stops and held back by an adjustable spring. On closing the circuit, the armature is brought sharply down upon the front stop, and on breaking the circuit, it is drawn back by the spring against the other stop. Thus each signal consists of two sharp clicks, separated by a longer or shorter interval of time. The short signals are termed dots and the longer ones dashes. The Morse code is made up of combinations of these dots and dashes. The armature is thus made to duplicate every motion of the sender's key. In the recording form of the in-

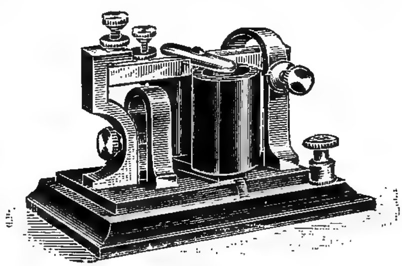


FIG. 174.

strument a fine style attached to the armature traces the signals upon a strip of paper actuated by a system of clockwork. In present practice all operators read these signals by sound.

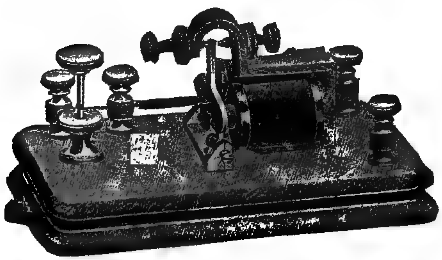


FIG. 175.

In long lines the resistance is frequently so great that the current becomes too weak to operate the sounder. In this case the armature, properly insulated, is connected to one terminal of a new circuit and the stop to the other. By this means the armature acts as a new sending key, closing and opening a new circuit. This device is termed a *relay* (Fig. 175). All sounders on the new circuit repeat the message strongly.

The actual arrangement of these instruments is shown in Fig. 176. The switches *s* and *s′* short circuiting the sending keys *k* and *k′* are closed, and a current is kept flowing through

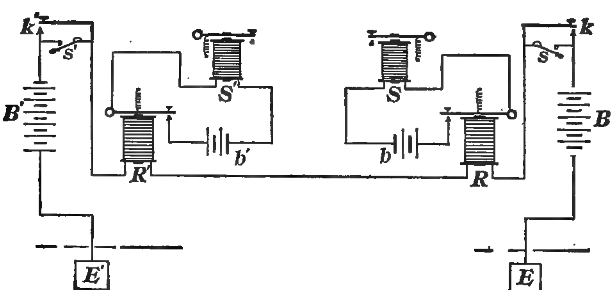


FIG. 176.

the circuit. If the operator on either side wishes to send a message, he opens his switch and operates his key, thus controlling his own sounder and all the instruments along the line.

322. Magnetization of Iron. Since the permeability of iron is very large in comparison with that of air, the number of lines of induction is greatly increased by the introduction of an iron

core into a solenoid. This furnishes an excellent method for studying the magnetic behavior of iron in fields of different intensity.

It is best to have the iron in the form of a ring, surrounded by a solenoid. In this case the iron fills the whole space in which a magnetic field exists (Art. 319), and no lines of induction pass through the air.

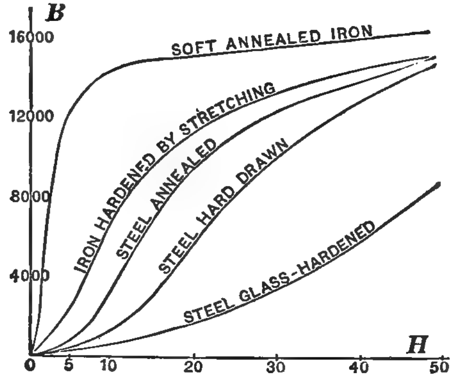


FIG. 177.

Let the field intensity be slowly increased from zero to larger values. The induction increases slowly at first, but soon rises rapidly, the rate of increase falling off again as still higher values of the magnetizing intensity are reached. On a piece of squared paper we may lay down the magnetizing field intensity H as abscissae and the induction B as ordinates, each referred to any convenient unit (Fig. 177). Curves of this kind are called *magnetization curves.*

The ratio μ between the induction B in the iron and the magnetizing field intensity H is called the *magnetic permeability of the iron* and defined by the equation

$$\mu = \frac{B}{H} \tag{370}$$

This quantity has its largest value at that point on each of the curves where a straight line from the origin touches the convex side of the curve.

In soft iron μ reaches values as high as 2000 or more. From the figure it is evident that the permeability of iron is not a constant, but varies with the intensity of the magnetizing field. The softer the iron, the steeper is the curve in the region of

small field intensities. Other ferromagnetic substances show similar curves, but their permeability is much smaller than that of soft iron.

323. Magnetic Hysteresis. The magnetization curve is obtained by subjecting an unmagnetized piece of the substance to the influence of a constantly increasing field intensity. But if the intensity be varied between two limiting values, $+H$ and $-H$, the curve assumes a different form (Fig. 178). Starting with a large intensity $+H$, a high value for B is obtained, corresponding to the point P. Upon decreasing H, B decreases also, but at a much slower rate than would be expected from the magnetization curve.

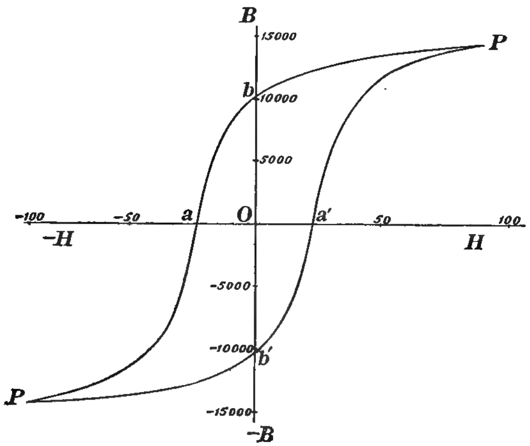


FIG. 178.

If H be reduced to zero, there is still a considerable induction in the iron, represented in the figure by Ob, which is called the *remanence*. If now the magnetizing field be reversed, by reversing the current in the solenoid, the induction falls off rapidly, and reaches a zero value with a small negative field intensity Oa, which is called the *coercive force*. Upon further increase of the negative field, the induction in the iron also becomes negative, and finally reaches a value, $-B$, equal in magnitude and direction but opposite in sign to that with which the experiment began, and represented on the curve by the point P'. A similar curve, but passing through the axes on the opposite side from the first curve, is obtained by returning to the original intensity, $+H$.

During a complete cycle of the field intensity the induction describes a loop, called the *hysteresis curve*,[1] showing plainly *hysteresis*, *i.e.* a *lagging* of the *induction* behind the *magnetizing field*. The area included by the curve is a measure of the loss of electromagnetic energy in the iron during the cycle. This energy appears as heat in the iron, and is lost for all practical purposes. It is therefore of great importance to use soft iron or mild steel in any electrical machine in which magnetization in a variable field occurs, since these metals have narrow hysteresis curves and small hysteresis losses.

324. Magnetic Flux. The magnetic state of a substance at a given point is characterized by the induction at that point (Art. 243). We shall now find it of advantage to restrict ourselves to no particular points in the body, but rather to consider the body as a whole. The *magnetic flux* through a given area is the total number of lines of induction passing through that area, or if we have to deal with a uniform field, in which the induction is the same for all points, the flux is the *product of the induction into the area.*

$$\Phi = B \cdot A \tag{371}$$

In general, the induction is not constant over the area, and the calculation of the flux requires the use of calculus.

325. Magnetomotive Force. In magnetism we may form a concept very similar to that of electromotive force in electricity. It is called *magnetomotive force*, and may be regarded as the cause of the magnetic flux. Following the same line of argument (Arts. 265 and 266), we may measure magnetomotive force by the work done in carrying a magnetic pole once around a complete magnetic circuit. Of course this cannot actually be done in the case where the magnetic field is due to a piece of magnetized iron. But the lines of induction produced by a current are closed lines passing through air only, so that the pole need not traverse a solid in making a complete circuit.

In the case of a ring solenoid, carrying a current of I amperes,

[1] *For experimental determination of magnetization and hysteresis curves, see Manual, Exercises 78 and 79.*

there exists inside *the spiral* forming the ring a uniform field of intensity

$$H=\frac{4\,\pi nI}{10}$$

or, substituting

$$n=\frac{N}{L} \tag{372}$$

where N is the total number of turns and L the length of the solenoid,

$$H=\frac{4\,\pi NI}{10\,L} \tag{373}$$

The force acting on a magnetic pole m at a point where the intensity is H is

$$F=Hm$$

and the work done in carrying the pole once around the circuit of length L is

$$W=FL=\frac{4\,\pi NI}{10}m \tag{374}$$

The ratio of the work done to the pole strength

$$\Omega=\frac{W}{m}=\frac{4\,\pi NI}{10}=1.257\,NI \tag{375}$$

is the magnetomotive force. It is numerically equal to the work done in carrying unit pole once around the magnetic circuit, and is independent of the cross section of the solenoid.

Since in the above equation N denotes the total number of turns, and the expression for Ω depends only upon N and I, the work done in carrying unit pole once around a single wire, carrying a current of I amperes, is $\frac{4\,\pi I}{10}$ ergs. This value is evidently independent of the path chosen.

The unit of magnetomotive force in the C. G. S. system is one erg per unit pole. A unit more frequently used is the *ampere turn*, or the magnetomotive force, produced by a single loop of wire, carrying a current of one ampere.

326. Law of the Magnetic Circuit. In the case of an iron ring surrounded by a magnetizing solenoid, the relation between

the magnetic flux and the magnetomotive force may be readily calculated, and also written in a form closely resembling Ohm's law. For we have the following equations for the flux:

$$\Phi = BA = \mu HA = \mu A \frac{4\pi N}{10L} I = \frac{4\pi NI}{10\frac{L}{\mu A}} = \frac{\Omega}{R} \tag{376}$$

where

$$R = \frac{L}{\mu A} \tag{377}$$

R is called the *magnetic reluctance* and is quite similar to electrical resistance, being proportional to the length of the circuit, inversely proportional to the area and inversely proportional to the permeability of the medium. Permeability thus corresponds to electrical conductivity.

When the magnetic circuit is not uniform, the reluctance of each part must be determined separately. The reluctance of the whole circuit is the sum of the reluctances of its parts. For example, the magnetomotive force of an electromagnet may be readily calculated from the number of its ampere turns. The air space between the poles introduces a reluctance very much larger than that of the iron. If now an armature be placed upon the magnet, the reluctance of the circuit is made quite small and the flux is greatly increased. The more closely the armature fits upon the magnet, the larger will be the total number of lines of induction, and consequently the greater will be the attractive force between magnet and armature.

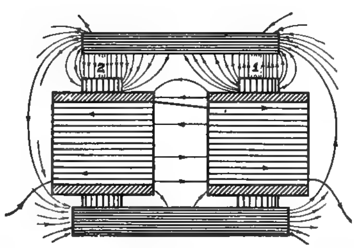


FIG. 179.

If a small air space be left between magnet and armature (Fig. 179), the circuit may be considered as consisting of four parts: the electromagnet, the air gap 1, the armature and the air gap 2. Assuming the induction to be uniform in each part, and distinguishing by subscripts the length, cross section and permeability of the different parts, the law of the magnetic

circuit becomes in this case

$$\Phi = \frac{1.257\,NI}{\frac{L_1}{\mu_1 A_1} + \frac{L_2}{\mu_2 A_2} + \frac{L_3}{\mu_3 A_3} + \frac{L_4}{\mu_4 A_4}} \qquad (378)$$

* **327. Magnetic Leakage.** Owing to their mutual repulsion, the tubes of induction have a tendency to spread out, especially in substances of small permeability, such as air. The characteristic figures for the magnetic field in the neighborhood of magnets, studied in Art. 244, show this clearly. For the same reason the tubes crossing the air space between the electromagnet and the armature (Fig. 179) are not parallel to each other. Some leave the electromagnet at the side, and may even pass to the other side without entering the armature at all; others enter the armature at the side without contributing to the attractive force. This spreading of the lines of induction is called *magnetic leakage*, and must be taken into account in all accurate calculations of the flux in a magnetic circuit.

CHAPTER XLI

ELECTROMAGNETIC INDUCTION

328. Induction by Magnets. If a magnet be thrust into a coil of wire whose ends are joined to a galvanometer, the galvanometer shows a momentary deflection, proving that, while the magnet is in motion with respect to the coil, a current flows through the coil, owing to the establishment of an E. M. F.

Upon removing the magnet from the coil, a deflection of the galvanometer in the opposite direction is observed. If the motion of the magnet be quite slow, the deflection is much smaller than if it be thrust in quickly. The same effects are observed if the magnet be kept stationary and the coil be thrust over the magnet.

These facts, which are known as the phenomena of *electromagnetic induction*, were discovered by Faraday[1] in 1831. They form the basis upon which all our electrical industries have been developed.

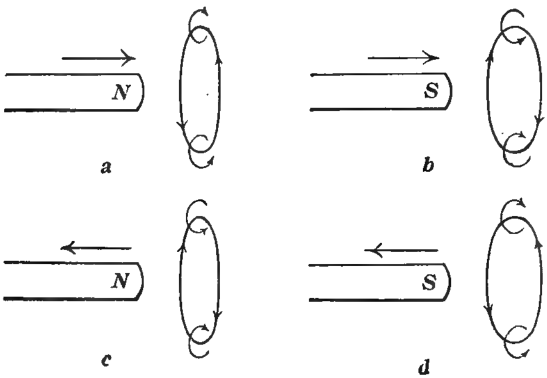
329. Lenz's Law.[2] The experiments of the preceding article show that the induced E. M. F. depends upon a change of position of a conductor in a magnetic field, or, in the language of Faraday, who considered a magnetic field to be filled with lines of induction, the E. M. F. is produced by a change in the number of lines of induction passing through the coil. *The E.M.F. induced is always such that its effect opposes the action which induces it.* This is known as Lenz's law.

Thus, if the relative motion of field and coil be such that the number of lines through the coil *increases*, the current produced in a closed coil will tend to weaken the magnetic field or to set up a field in the *opposite* direction; if the motion be such as to *decrease* the number of lines, the current will set up a field in the *same* direction, or tend to strengthen the existing field. From this it follows that the induced current produces

[1] Faraday, *Experimental Researches, Series I, Phil. Trans.*, 1831.

[2] E. Lenz, *Pogg, Ann.* 31, p. 483, 1834.

a north polarity at the side towards an approaching north pole (Fig. 180 *a*), and a south polarity facing an approaching south pole (Fig. 180 *b*). This induced current flows in the opposite sense while the pole is being withdrawn (Figs. 180 *c* and 180 *d*).



a b c d

FIG. 180.

The system, consisting of a magnetic field and a closed conductor, acts as if there were a certain inertia in the field, or a tendency to maintain, unchanged, the number of lines threading through the circuit. It requires an expenditure of energy to produce a *change in the configuration of the system.* This property of the electromagnetic field may be called *electromagnetic inertia.*

The energy expended to overcome this inertia may be measured in two different ways: (*a*) by the mechanical work necessary to move the magnet; that is, by the product of the force applied into the distance through which the force acts; (*b*) by its counterpart, the electrical energy which appears in the conductor. This is measured by the product of the induced E. M. F. into the quantity of electricity flowing through the circuit.

It requires, therefore, the application of a mechanical force to overcome the opposition to a change of configuration in the electromagnetic system. As soon as the motion of the magnet is stopped, the induced current ceases to flow. We have thus added to the methods previously described a third method for generating an electric current, namely, by transforming mechanical energy into electrical energy. All modern generators, designed to give large electric currents, are constructed upon this principle.

330. Magnitude of Induced Electrical Quantities. We have seen that the deflection of a galvanometer, placed in a circuit in which a current is produced by electromagnetic action, depends upon the time rate at which the magnetic field through the coil changes. Quantitative experiments have shown (a) that for a given coil *the induced* E.M.F. *is proportional to the time rate of change of the number of lines of induction through the circuit;* (b) that for the same rate of change of lines of induction *the induced* E.M.F. *is proportional to the number of turns in the coil.* If the E.M.F. be expressed in C.G.S. units, these relations are given by the equation

$$E = -N\frac{\Phi_2 - \Phi_1}{t} = -N\frac{d\Phi}{dt} \tag{379}$$

where the negative sign indicates that the E.M.F. is opposed to the action producing it.

The E.M.F. is variable unless the time rate of change of the magnetic flux be constant. The expression $-N\frac{\Phi_2 - \Phi_1}{t}$ denotes in any case the *average* E.M.F. *during the time* t.

If it be desired to express the E.M.F. in *volts*, the equation takes the form

$$E = -\frac{N}{10^8}\frac{\Phi_2 - \Phi_1}{t} = -\frac{N}{10^8}\frac{d\Phi}{dt} \text{ volts} \tag{380}$$

since the volt is equal to 10^8 C.G.S. units.

The induced E.M.F. is restricted to that part of the circuit where a change of magnetic flux occurs. Those parts of the circuit which do not *cut* lines of induction, do not contribute to the induced E.M.F. in any way. Induction takes place whether there be a closed circuit or not. Of course no current is produced on open circuit, but an E.M.F. is always produced by a change in the magnetic flux through a coil. If the coil be closed and have a resistance of R ohms, the average current passing through the circuit during the time t is

$$I = -\frac{N}{10^8 R}\frac{\Phi_2 - \Phi_1}{t} \text{ amperes} \tag{381}$$

and the total quantity of electricity passing through the circuit during the time t in which the flux changes from Φ_1 to Φ_2 is

$$Q = It = -\frac{N}{10^8 R}(\Phi_2 - \Phi_1) \text{ coulombs} \qquad (382)$$

The total quantity is therefore *independent of the time* and is proportional to the total change in magnetic flux through the coil.

331. Induction by Currents. It is, of course, immaterial how the magnetic field which passes through the circuit is produced. Instead of the magnet, used in Art. 328, we might just as well have used a solenoid. This solenoid circuit is called the *primary*, while the circuit in which the E.M.F. is induced is called the *secondary circuit.* If the primary be brought toward the secondary, Lenz's law requires that the induced current flow in the opposite sense to the inducing current. When the primary is moved away, both currents must flow in the same sense.

The experiment may be varied in the following manner: Let the two coils, the primary being on open circuit, be placed in a fixed position close to each other. If now the primary be closed, the magnetic effect of its current is evidently the same as if the closed primary, carrying with it its full number of lines of induction, had approached from infinity to its fixed position. If the current of the primary be opened, the effect is the same as if the primary, with its magnetic field, had been removed to infinity. Thus the *make* of the current in the primary induces a secondary current in the opposite sense to that of the primary current, while the *break* of the primary induces a current in the same sense.

The best effect is obtained when the primary coil is placed inside the secondary coil, since in this position all lines of induction from the primary pass through the secondary.

It is to be noted that in all cases the induced current lasts only so long as the magnetic field of the primary is *changing*, and disappears as soon as the primary current reaches a constant value.

If, with a steady current through the primary, a soft iron core be thrust inside the primary coil, a violent deflection of

the galvanometer indicates a large induced E.M.F. in the secondary circuit. It is clear that this has been due *not* so much to any change in the primary *current* as to a change in the *magnetic flux* due to the high permeability of the iron core. For the same reason correspondingly large throws of the galvanometer will now be observed on opening or closing the primary circuit.

332. Mutual Inductance. If a current be sent through the primary coil of cross section A, a magnetic flux equal to μHA is produced. This flux is proportional to the current and number of turns in the primary. It is also dependent upon the dimensions of the primary coil. A definite portion Φ, of this flux also passes through the secondary of the coil, the amount depending upon the relative position of the two coils. If there be N turns in the secondary coil, and if we call $N\Phi$ the *coil flux* through secondary, then this coil flux is proportional to the current in the primary. It is also dependent upon the *dimensions*, the *number of turns in both coils* and upon *their relative position.* Now for two definite coils in a definite position, all iron being excluded, these last three factors *are all constant.* We may therefore write for the coil flux through the secondary,

$$N\Phi = MI_1 \tag{383}$$

where I_1 is the current through the primary and M is a constant depending upon the dimensions, number of turns and relative position of the two coils. This constant is called the *coefficient of mutual induction*, or the *mutual inductance* of the two coils. Since M is a constant, the variations of the coil flux depend only upon the variations of the current. The induced E.M.F. in the secondary E_2, is therefore

$$E_2 = -N\frac{d\Phi}{dt} = -M\frac{dI_1}{dt} \tag{384}$$

The *mutual inductance* of two coils is therefore *the ratio of the E.M.F. induced in one of the coils to the time rate of change of current in the other.*

333. Self-inductance. If in a circuit consisting of a battery, a coil of many turns and an interrupter or key, the current be

repeatedly made and broken, it will be seen that a bright spark appears *at each break* of the current. This spark prevents the current from falling at once to zero. Faraday called this phenomenon the "extra current." Similarly, on closing the circuit, the current does not instantly assume its full value, as indicated by Ohm's law, but rises more or less rapidly to this maximum value. Both these phenomena are more marked in coils of many turns, or in coils containing an iron core. Thus, in a large electromagnet, it may take a *number of seconds*, or *even minutes*, for the current to assume its maximum value.

Both the phenomenon of the "extra current" and that of the gradual rise of the current in a circuit are easily explained by the electromagnetic action of the coil upon itself, and these phenomena constitute what is called *self-induction*. When the current is closed, a counter E.M.F. is set up in the coil opposing the establishment of the current; when the circuit is opened, the induced E.M.F. tends to continue the current, and thus produces the spark at the gap.

That such a result is to be expected is evident when we remember that a magnetic field represents a certain amount of energy, and that this energy must be supplied from the energy of the current itself during the building up of the field. During this short period, therefore, a part of the energy will not appear as energy of current in the wire, but as magnetic energy in the field, and the current will consequently be smaller during this time than after the field has been established. Here, again, the electromagnetic field shows a property very similar to that of inertia.

Since self-induction is only a special case of induction, equation (379) must hold. The coil flux is proportional to the current in the circuit, and otherwise depends only upon the form and number of the turns in the given coil. We have, therefore, in this case,

$$N\Phi = LI \tag{385}$$

and

$$E = -N\frac{d\Phi}{dt} = -L\frac{dI}{dt} \tag{386}$$

The constant L is called the *coefficient of self-induction*, or the *self-inductance* of the coil, and may be defined as *the ratio of the induced counter* E.M.F. *to the time rate of change of current in the coil.*

334. Energy stored in the Field. The effect of self-inductance may be shown by making and breaking the current in a circuit containing a large electromagnet and an incandescent lamp in parallel (Fig. 181). The resistance in the circuit should be so adjusted that with a steady current the lamp L will be only dull red. When the circuit is closed through the key k, the E.M.F. induced in the coil of the electromagnet opposes the flow of electricity through the coil, and its effect is the same as if a high resistance were temporarily inserted in the inductive branch. Consequently, since the lamp is practically non-inductive, the current through it will at first be larger than after the current has become constant, and the lamp will light up for a moment, owing to the current flowing through it from C to D. Upon breaking the current, the lamp again flashes up, since the energy stored in the magnetic field now reappears as the energy of a current passing through the lamp from D to C. In this case the electromagnet becomes for an instant a generator of an electric current.

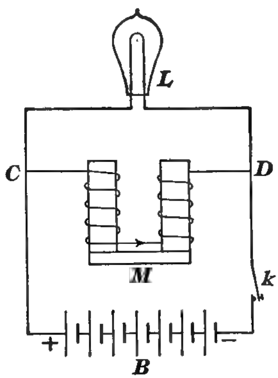


FIG. 181.

Let us assume that in a circuit of self-inductance L the current rises uniformly in the time t from zero to I amperes. The average current during this time is $\frac{1}{2}I$, and the quantity of electricity flowing through the circuit is

$$Q = \tfrac{1}{2}It \tag{387}$$

The rate of change of the current being uniform, the induced E.M.F. is constant during the time t, and

$$E = \frac{LI}{t} \tag{388}$$

The quantity of electricity Q is forced through the circuit against E, and the work done by the electric current in time t against the opposing E. M. F. E is

$$W = EQ = \tfrac{1}{2} LI^2 \tag{389}$$

The same result may be obtained by the use of calculus when the rate of change of current is not uniform. The above expression for the energy stored in the magnetic field depends, therefore, only upon the self-inductance of the circuit and upon the current passing through it.

335. Unit of Inductance.[1] Mutual and self-inductance are physical quantities of the same nature, and the *same unit* must be used *for both.* The unit of inductance is the inductance in a circuit in which the induced E. M. F. is one volt when the inducing current changes at the rate of one ampere per second. This unit is called the *henry*, after the American physicist, Joseph Henry (1799–1878). It is equal to 10^9 C. G. S. units.

336. The Induction Coil. An induction coil (Fig. 182) consists of a primary coil PP', of relatively few turns and low resistance, surrounding a core of soft iron wires. A secondary SS', consisting of many turns of fine wire, is wound around the primary. When a current is made or broken in the primary, the E. M. F. in the secondary becomes sufficiently high to cause a spark across the terminals of the secondary. Since the *break* of the current is always much more sudden than the *make*, the former produces the higher E. M. F.

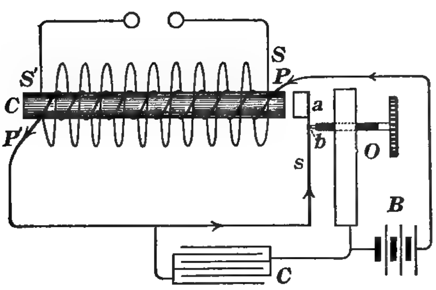


FIG. 182.

In large induction coils the spark may be made as long as 50 cm or more. Since the sparking potential in air between

[1] *For experimental determination of inductances, see Manual, Exercises 74 and 75.*

spheres of 1 cm diameter, separated by a distance of 1 cm, is 27,000 volts, we see that the E. M. F. between the terminals of large coils may reach a value of more than a million volts. The quantity of electricity flowing in the secondary is, however, quite small, since the energy produced in the secondary can never be equal to that expended in the primary.

In order to obtain such high electromotive forces, it is clear from equation (379),

$$E = -N\frac{d\Phi}{dt}$$

that this end may be attained by increasing each of the factors N and $\frac{d\Phi}{dt}$. This suggests that N, the number of turns in the secondary coil, must be very large; in some cases it reaches many hundred thousand. Again, the factor $\frac{d\Phi}{dt}$ may be increased both by making $d\Phi$ large and dt very small. The change in the magnetic flux $d\Phi$ may be increased by using a large current in the primary and by using iron of high permeability in the core. The value of dt may be greatly reduced by using some form of interrupter whereby the current may be made and broken with great rapidity. This rapid make and break of the primary current is usually effected by means of some form of automatic interrupter acting on the principle of the hammer in the electric bell (Art. 320).

337. Action of the Condenser. Owing to the large self-inductance of the primary circuit, a spark tends to jump across the gap b at each break, and thus to interfere with the sudden interruption of the current. This spark would not only burn off the contact point and ruin the automatic break, but it would also prolong unduly the fall of the current to zero. In order to suppress this spark, a condenser C (Fig. 182) is placed in parallel with the spark gap. This serves a double purpose. First, the extra current at the break, instead of jumping the gap in the form of a spark, is diverted into the condenser and charges it to a high difference of potential. By this means the spark

and the burning of the contact are both avoided, and the time of break dt is greatly reduced. Second, the condenser, immediately after the break and before the next make of the current, discharges backward through the battery and the primary coil, thus sweeping out all lines of induction remaining in the core, and inserting others in the opposite direction. The value of $d\Phi$ is thus greatly increased.

When the current through the primary is closed, the self-inductance of the coil retards the rise of the current. For this reason only the break of the current produces a sufficiently high E. M. F. in the secondary to produce a spark over a wide gap, while that due to the make is unable to do so. The spark in the secondary of an induction coil is therefore, in general, unidirectional and due only to the break in the primary. Hence it is customary to speak of the positive and negative terminals of an induction coil.

*** 338. The Wehnelt Interrupter.** The Wehnelt interrupter is a very efficient device for breaking the primary circuit. It consists of an electrolytic cell containing a large lead cathode and a very small platinum anode, usually a short piece of fine platinum wire, projecting through a hole in a glass or porcelain tube into the electrolyte, which is dilute sulphuric acid. The large current density at the wire produces sufficient heat to form a layer of vapor around the anode, and interrupts the current. The extra current throws off the gas film and reëstablishes electric connection. Thus the current is alternately made and broken. The number of interruptions may be made as high as 2000 per second. No condenser is needed when a Wehnelt interrupter is used. In this case the passage of electricity between the terminals of the secondary appears more like a continuous stream than a succession of separate sparks.

339. Eddy Currents. If a mass of copper be placed near a swinging magnet, the magnet will be brought to rest much more rapidly than if the copper were absent. The reason for this is that the moving magnetic field, sweeping through the conductor, produces in it currents by induction, which in front

of a moving pole, present a pole of the same sign; behind it one of the opposite sign (Fig. 183). These currents are called *eddy currents*, or *Foucault currents*. They give rise to a mechanical force tending to stop the swinging of the magnet.

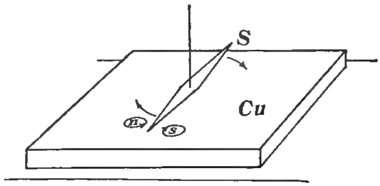


FIG. 183.

Again, when a plate of metal, preferably of copper, swings like a pendulum between the poles of an electromagnet, the swinging motion will immediately be stopped as soon as the electromagnet is excited. If, however, slits be cut in the plate preventing the establishment of eddy currents, the effect of the magnetic field is greatly diminished. The same principle is applied in the magnetic brake, where a thick disk of copper attached to a rotating axle is brought between one or more pairs of poles of an electromagnet. On exciting the magnet, the axle is promptly brought to rest.

The magnetization of a solid piece of iron, placed inside a solenoid, is much retarded by eddy currents which flow in such a direction as to oppose the magnetization. Therefore, where rapid magnetization or demagnetization of an iron core is desired, as in an induction coil, the core should consist of laminated iron or of a bundle of iron wires, insulated from each other.

The establishment of eddy currents always means a loss of energy, since their energy is transformed into heat in the conductor and is of no further practical use.

340. The Telephone. The telephone is an important application of electromagnetic induction. The telephone line, in its simplest form, consists of a metallic circuit in which are inserted at two stations, *A* and *B* (Fig. 184), two bobbins of fine wire, each encircling the end of a small, permanent bar magnet. Close to the end of each bar magnet is placed a thin iron disk

FIG. 184.

or diaphragm which, by its vibrations, may be made to approach or recede from the bar magnet through a small distance.

When a sound is produced in front of either of these diaphragms, it vibrates in unison with the sound vibrations. As the disk approaches the magnet, the reluctance of the magnetic circuit is reduced, the flux through the magnet is increased, and consequently, by induction, a current is set up in the electric circuit. The induced current will be in the opposite direction when the disk moves away from the magnet. At the receiving station the current, in flowing through the bobbin, alternately strengthens and weakens its magnet, producing a vibration of the diaphragm and of the air in front of it in exact coincidence with the vibrations of the diaphragm at the sending station. In this manner the sound waves at the first station are exactly reproduced at the second.

It will be noticed that no battery is needed in this arrangement. The distance through which transmission of speech is possible in such a simple system is limited, because the currents thus produced are quite feeble.

341. The Transmitter. In modern practice the instrument described in the last article is used as a receiver only. The *transmitter* (Fig. 185) consists of a diaphragm D, behind which is a small chamber g, filled with granular carbon. The current from a battery, usually of dry cells, is sent through this carbon resistance. As the diaphragm vibrates in response to a sound, it changes the pressure between the carbon particles. Since the resistance of loose contacts varies considerably with pressure, the vibrations of the diaphragm produce relatively large variations of the current, in unison with the sound waves.

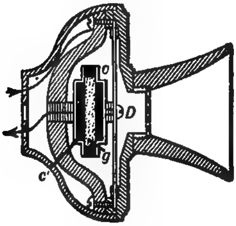


FIG. 185.

These currents may be sent directly through the circuit to the receiver at the other station. More frequently, however, they pass through the primary of a small induction coil, while the currents induced in the secondary travel over the line and act upon the receiver in the manner described in Art. 340.

***342. The Modern Telephone Service.** Fig. 186 illustrates the manner in which, in the system of the Bell Telephone Company, a subscriber's line is connected to the central station. The battery *B* is no longer at the subscriber's end, but at the central station. No current flows through the line as long as the receiver *R* hangs upon its hook, since the condenser *C* prevents a passage of electricity through the electromagnets, *e*, of the bell. When the receiver is taken off, contact is made at the point *a*, and a current flows through the transmitter *T*. At the same time a small electromagnet *e′* at the central station is excited by the current, and attracts an armature, closing a shunt circuit, and lighting a small incandescent lamp *L*, placed at the number of the subscriber calling up central. The glow of the lamp attracts the attention of the operator, who

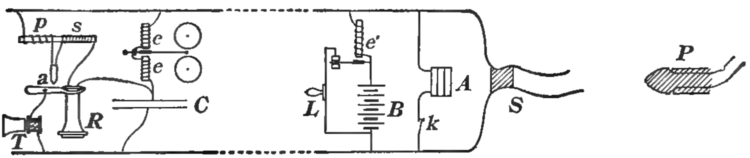


FIG. 186.

connects by a switch her own receiver to the subscriber's line and ascertains the number which is wanted. She then makes connection between the two lines by means of flexible, conducting cords, with plugs *P* at their ends. These are pushed into the sockets which form the terminals of the lines of each subscriber. In order to call up the desired number, the operator connects this line to a source of alternating current *A*. The condenser *C* will be charged alternately, first in one and then in the other direction, and an alternating current surges through the electromagnet *ee* and rings the bell.

As soon as the subscriber at the second station removes the receiver from the hook, the circuit through the transmitters in both stations is completed, and communication between them is rendered possible. The receiver is connected to a secondary coil *s*, but in this system this coil is in series with the primary. However, it has been found that such an arrangement works as well as if the primary and secondary coils were separated.

CHAPTER XLII

DYNAMO-ELECTRIC MACHINES

343. The Dynamo. A dynamo-electric machine, or a dynamo, is a machine in which either mechanical energy is transformed into electrical energy, or electrical energy into mechanical energy. Precisely the same form of construction is used for both purposes.

In the first case, a mechanical force applied to a conductor moves it across a magnetic field and generates in it an induced E. M. F. An electric current will flow through an external circuit connected to the moving conductor. A dynamo used in this manner is called an *electric generator* (Arts. 344 to 352). In the second case, a current is sent through a conductor placed in a magnetic field. This produces a mechanical force, which, acting upon the conductor, causes it to move. A dynamo used for the production of mechanical motion is called an *electric motor* (Arts. 353 to 358).

344. The Generator Rule. If a straight wire be moved across the lines of induction of a magnetic field, the reaction due to the induction tends to oppose the motion (Art. 329). Using Faraday's mode of expression, the induced current will flow in such a direction that the lines of induction are crowded together in front of the moving wire, resisting the displacement. Let AB (Fig. 187) represent an element of the circuit parallel to the z-axis in a field F, whose lines of induction are parallel to the positive y-axis. Let the motion m of the conductor be parallel to the x-axis towards the observer.

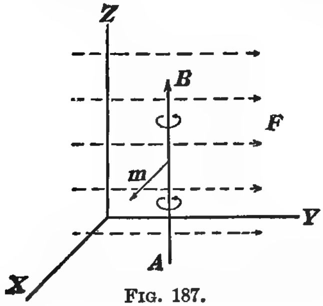


FIG. 187.

The induced current will then flow in such a direction that the field in front of the moving conductor is intensified, or lines of induction are added by the current in the same sense as the original field, while behind the conductor the lines due to the current are oppositely directed, and weaken the field. According to the right-hand rule (Art. 256), the current will flow upward from A to B, or in the positive direction of the z-axis. The relation between the three directions may be remembered by the following rule: The *motion* in a magnetic *field* produces a *current* in such a direction that these three quantities form a right-handed coördinate system, if taken in the above order.

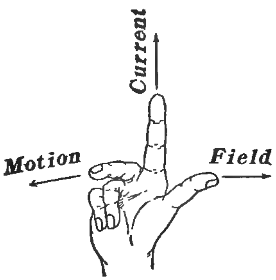


FIG. 188.

If the thumb, index and middle finger of the *right* hand be held at right angles to each other (Fig. 188), the thumb (first finger) F points in the direction of the field, the index finger I in the direction of the current, and the middle finger M in the direction of the motion.

345. Quantitative Relations for Generator. Let a wire AB (Fig. 189) slide sidewise along two straight wires, CE and DF, which are l cm distant from each other, and together with AB form a closed circuit. If there exist a uniform magnetic field perpendicular to the plane of the circuit, and if AB move with a uniform velocity v across this field, the area covered by the wire in t seconds is lvt cm², and the total number of lines, cut in t seconds, is

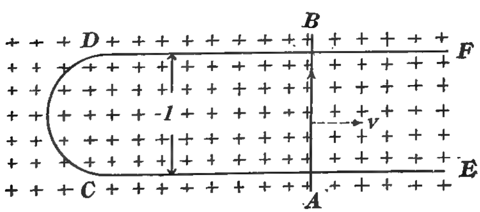


FIG. 189.

$$\Phi = Blvt \tag{390}$$

The E. M. F. induced in the sliding wire AB is equal to the rate at which the lines are cut (Art. 330). Therefore, disregarding the sign,

$$E = \frac{Blvt}{t} = Blv \text{ C. G. S. units} = \frac{Blv}{10^8} \text{ volts} \qquad (391)$$

The direction of this E. M. F. is determined by the generator rule. Thus, if the field (Fig. 189) be directed into the paper, and the conductor move from left to right, then the E. M. F. will be directed from A to B.

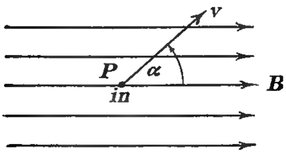


FIG. 190.

If the motion be not at right angles to the field, the above equation must be modified. For, let the wire be at right angles to the plane of the paper, crossing it at the point P (Fig. 190), and let B be parallel to the plane of the paper, that is, at right angles to the wire. If now the wire be moved in a direction v, making an angle α with B, the number of lines of induction, cut in t seconds, is

$$\Phi = Blvt \sin \alpha \qquad (392)$$

and $$E = Blv \sin \alpha \text{ C. G. S. units} = \frac{Blv}{10^8} \sin \alpha \text{ volts} \qquad (393)$$

and the E. M. F. is directed *into* the paper.

Obviously no E. M. F. is induced if the wire be moved parallel to the lines of induction.

346. Faraday's Disk. The first electric generator was constructed by Faraday, who rotated a copper disk between the poles of a magnet (Fig. 191). Each radius of the disk cuts the lines of induction at right angles, and thus becomes the seat of an induced E. M. F. If each radius sweep out an area a in time t, and if A denote the total area of the disk swept out in time T, then for n uniform revolutions per second we have

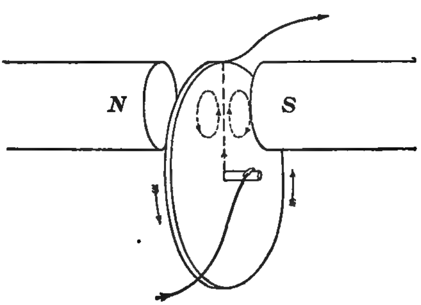


FIG. 191.

$$\frac{a}{A} = \frac{t}{T} \qquad (394)$$

The flux through area a is, for a uniform field,

$$\Phi = aB = AB\frac{t}{T} = AnBt \tag{395}$$

and the induced E. M. F. is

$$E = \frac{\Phi}{t} = AnB = \pi r^2 nB \text{ C. G. S. units} = \frac{\pi r^2 nB}{10^8} \text{ volts} \tag{396}$$

where r is the radius of the disk.

If an electric circuit, containing a galvanometer, be connected to the axle and to the circumference of the disk, the current produced by this machine will flow as indicated in the figure.

347. A Loop of Wire rotating in a Magnetic Field. Another simple electric generator consists of a plane rectangular loop (Fig. 192) rotating with uniform angular velocity around its longer axis. This axis of rotation is placed at right angles to a uniform magnetic field. From the previous discussion (Art. 345) it is clear that an E. M. F. is induced only in those wires each of length l, which cut the lines of induction. The two wires forming the ends of the loop may therefore be neglected, since they move at all times parallel to the lines of induction.

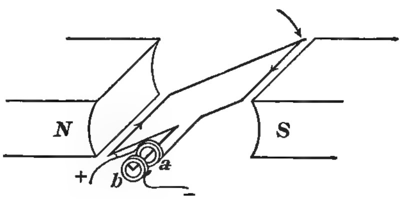


FIG. 192.

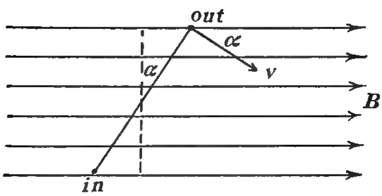


FIG. 193.

Let the plane of the coil at any instant make an angle α with the plane at right angles to the field (Fig. 193). Then, at the given instant, the two effective wires whose cross sections are shown at the points marked "out" and "in" are moving in a direction which makes an angle with the lines of induction. Their velocity at right angles to the field is therefore $v \sin \alpha$, and the E. M. F. induced in each wire is

$$E' = \frac{Blv \sin \alpha}{10^8} \text{ volts} \tag{397}$$

Applying the generator rule, it is easily seen that the two E. M. F.'s are in opposite sense with respect to the plane of the paper, but in the same sense in the electric circuit from one terminal of the loop to the other (Figs. 192 and 193), and that they must therefore be added to obtain the total E. M. F. of the generator, which is

$$E = \frac{2\,Blv \sin \alpha}{10^8} \text{ volts} \tag{398}$$

If now the terminals a and b (Fig. 192) of the loop be connected to two metal rings upon which two metal springs or brushes rub, then the E. M. F. induced in the loop will send a current through an external circuit attached to the two brushes. These brushes are called the terminals of the machine.

348. The Alternating Current. During one complete revolution of the rectangle, described in Art. 347, the induced E.M.F. increases from zero, when the plane of the coil stands at right angles to the field, to a maximum value of $2\,Blv \cdot 10^{-8}$ volts when its plane is parallel to the field; it then decreases to zero, reverses its sense, increases to $-\,2\,Blv \cdot 10^{-8}$ volts, and returns to its zero value after a complete revolution. If the E. M. F. be plotted as a function of the angle α, a sine curve is obtained (Fig. 194).

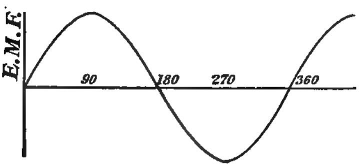


FIG. 194.

Such an E. M. F. is called an alternating E. M. F., and since the current in the circuit is proportional to the E. M. F., the resulting current is represented by a curve of the same general form as that of the E. M. F.

349. The Alternator. Machines producing an alternating E. M. F. are called alternators. The magnetic field is produced by a powerful electromagnet called the *field magnet.* Instead of a single loop of wire, a large number of turns are

used for the rotating part, and in order to make the magnetic flux through the rotating coils as large as possible, they are wound on laminated, soft iron cores (Art. 339). The rotating part, consisting of the coil, with its core, is called the *armature.*

The common form of alternator is always multipolar (Fig. 195). The winding of the field magnets is such that the polarity of adjacent poles is always of opposite sign. The E. M. F. induced in coils passing beneath a north pole is then in an

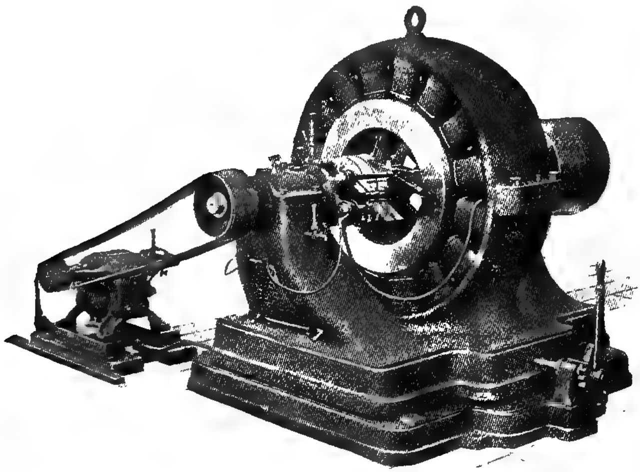


FIG. 195.

opposite sense to that induced in coils passing beneath a south pole. But since the direction of the armature winding changes between each two poles, the E. M. F.'s of all coils are in the same sense through the armature, and their effects are added. The direction of the current given by the machine changes when the coils pass the point midway between two adjacent poles.

The reason for using multipolar machines is that, for purposes of illumination, frequencies above fifty alternations per second are needed to prevent unpleasant flickering, and such speeds would be difficult to obtain with large bipolar machines.

350. The Transformer. If an iron ring (Fig. 196) be wound, as indicated, with two separate coils of insulated wire, P and S, and an alternating current be sent through the primary circuit P, it will be found that an alternating current of the same frequency flows through the secondary circuit S when this is closed. In this case, energy from the primary circuit has been transmitted to the secondary circuit through the *medium of the iron core.* If E_1, I_1 and E_2, I_2 be the electromotive forces and currents in the primary and secondary circuits respectively, then, neglecting the small losses due to hysteresis and eddy currents in the iron, we have for any small time interval dt,

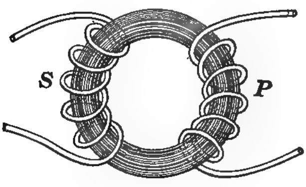


FIG. 196.

$$E_1 I_1 dt = E_2 I_2 dt \tag{399}$$

or

$$E_1 I_1 = E_2 I_2 \tag{400}$$

It may also be shown that the ratio between E_1 and E_2 is very nearly equal to the ratio between the number of turns of wire in the primary and secondary coils. In other words, a large alternating current of low electromotive force may be transformed into a small current of high electromotive force, or *vice versa*, through a proper choice of the number of turns in the two coils. Such a device is called a *transformer.*

In the commercial transformer (Figs. 197, 198) the core is made up of many thin sheets of soft iron or mild steel closely packed together. This form of laminated core is adopted to avoid eddy currents. Transformers are designated as *step-up* or *step-down* transformers, according as they are used to increase or decrease the voltage. In electric-lighting circuits the transformers are usually step-down transformers Thus, an E.M.F. of 1100 volts is not uncommon on electric-lighting mains. This would be dangerous for use in dwellings, so the voltage is reduced to 110 volts. To this end there are 10 times as many turns on the primary as on the secondary, which is connected to the circuit in the house. The efficiency of a good transformer is somewhere between 95 and 97 per cent.

It is of great advantage to use high voltages for the transmission of electric power, since in this way the energy loss due to heating is appreciably reduced. For example, if 10,000 watts be transmitted over the same line, in one case by an E.M.F. of 100 volts, in another case by one of 1000 volts, the currents would be 100 amperes and 10 amperes respectively. Since the heating effect is proportional to the square of the current, the heat loss in the first case would be 100 times larger than that in

FIG. 197.

FIG. 198.

the second. If the same loss be allowed, it is evident that the size of the conductor may be made much smaller when high voltages are used, and this means great economy in the construction of transmission lines. An upper limit to the voltage is set only by the difficulty of insulation. Voltages as high as 30,000 to 60,000 volts are not unusual in modern power transmission. In such cases the coils of the transformers are immersed in oil of high insulating power.

***351. The Polyphase Generators.** In polyphase current machines the armature consists of two or three separate coils. Fig. 199 represents the simplest form of a two-phase generator. The two coils on the armature are at right angles to each

other. When the E. M. F. in one coil is at its maximum, that in the other is zero, and as the armature rotates, the E. M. F. of one circuit is always 90 degrees ahead of that in the other. The two are said to differ in phase by 90 degrees. In the three-phase machine three separate coils are placed upon the armature in such a position that the phase difference between the E. M. F.'s is 120 degrees in each case.

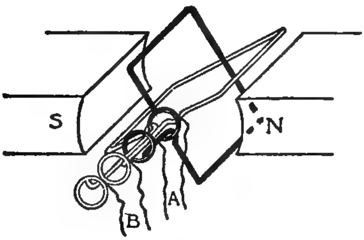


FIG. 199.

These machines are called polyphase machines in order to distinguish them from a machine giving but one alternating current, which is sometimes called a single-phase machine.

352. The Direct Current Dynamo. In order to obtain a current which is constant in direction, a *commutator* is used instead of the collector rings. For example, if there be but a single coil on the armature, the ring is split into two parts, which are insulated from each other (Fig. 200). The brushes sliding on the commutator are placed in such a position that they exchange contact between the two halves of the commutator when the current in the coil passes through zero and changes direction. In this manner one of the brushes is always kept at the higher and the other at the lower potential. The E. M. F. through the external circuit is all the time in one direction (Fig. 201), but it is pulsating, varying between zero and the maximum twice during each revolution of the coil.

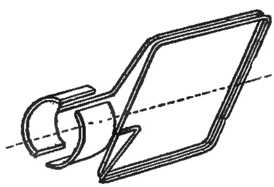


FIG. 200.

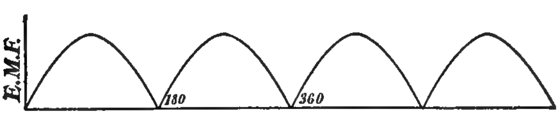


FIG. 201.

If two coils be used, placed at right angles to each other, a four-part commutator is needed, and the resulting E. M. F. may be considered as a steady, direct potential difference

Oa (Fig. 202), upon which are superposed during each revolution four pulsations of relatively small amplitude. In modern machines there are often many hundred coils, all connected in series, and placed in specially designed grooves, uniformly distributed over the armature. Every second or third coil is connected to one of the many sections of the commutator, thus making the E. M. F. practically constant.

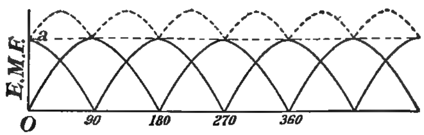


FIG. 202.

The current is taken off by the brushes from opposite sections of the commutator, which are connected to those turns of the armature in which the induced E. M. F. just passes through the zero value. The E. M. F. of the machine equals the sum of all the E. M. F.'s produced in the various coils between the brushes, and the resultant current is practically constant in strength. In direct current dynamos the E. M. F. can never be made as high as in an alternator, since the insulation between the sections of the commutator is not sufficient to support a difference of potential much higher than 500 volts.

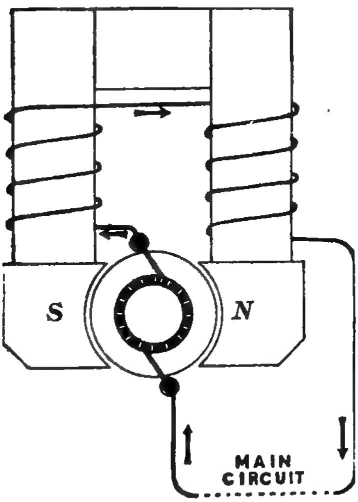


FIG. 203.

The field magnet is usually excited by means of a current taken from the machine itself. This may be accomplished in any one of three different ways. In the series-wound machine (Fig. 203) the field coils are in series with the external circuit. In the shunt-wound machine (Fig. 204) the field coils and the external circuit are in parallel, while the compound-wound machine is simply the shunt machine to which a few coils in series with the external circuit have been added. Each of these ma-

chines has been developed to meet special requirements as they have arisen in commercial practice.

It is also to be noted that although the cores of the field magnets are made of the best soft iron or mild steel, yet when the machine is stopped there is sufficient *remanence* in the soft cores to generate a weak current when the machine is started again. This weak current, circulating in the field coils, strengthens the field, and so induces a larger current, until current and field mutually build up to the maximum magnetization and maximum current which are possible under the circumstances.

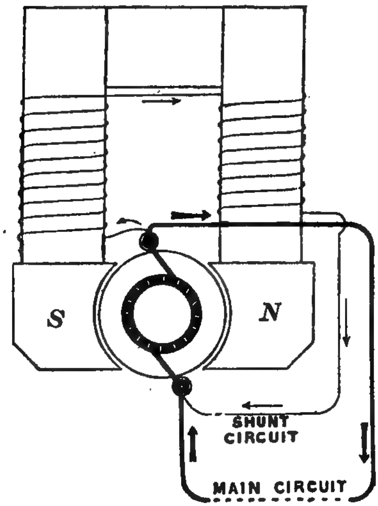


FIG. 204.

In new machines, when started for the first time, it is usually necessary to excite the field from some external source of current, although in many cases even this is unnecessary, the cores having already gained polarity owing to the hammering of the metal while in the earth's magnetic field.

353. Force upon a Conductor carrying a Current in a Magnetic Field. A conductor carrying a current, when placed in a magnetic field, is acted upon by a mechanical force. If, for example, a part of an electric circuit consist of a copper wire hung from a hook and dipping into a cup of mercury (Fig. 205) which surrounds a magnetic pole, the wire will begin to rotate around the pole as soon as the circuit is closed. The direction of rotation depends upon the relative directions of

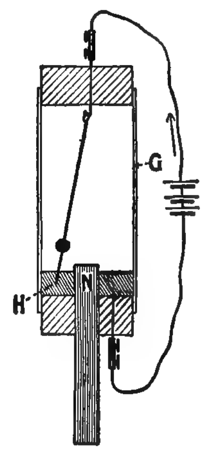


FIG. 205.

the current and of the field. Thus, if the pole be a north-seeking pole and the current flow towards the pole, the rotation will appear clockwise to a person looking down upon the pole. If the current through the wire be reversed, the rotation will be in the opposite direction.

Barlow's wheel (Fig. 206) is a metallic disk free to rotate about its center, and with its lower rim dipping into a trough of mercury between the arms of a permanent horseshoe magnet. When a current passes along the radius of the disk between the axle and the mercury, the wheel begins to rotate, and there is a transformation of electrical energy into mechanical energy. Barlow's wheel is a motor and the exact analogue of Faraday's disk (Art. 346). Both pieces of apparatus are of the same construction, and only the mode of using them determines whether we have a generator or a motor.

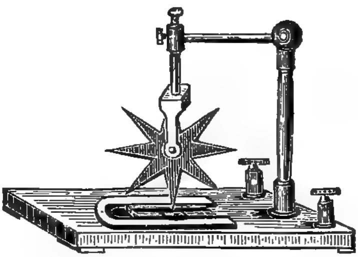


FIG. 206.

354. The Motor Rule. Consider an electric circuit of the same form as in Fig. 189 (Art. 345), but instead of applying a force to move the cross wire AB, let an electric current be sent through it. To fix our ideas, let the magnetic field pass downward into the paper, and the current flow from B to A. The right-hand rule (Art. 256) shows that the current through the wire produces lines of induction, entering the loop from above or in the same direction as the magnetic field. Thus the *number of lines* from both sources are *crowded inside* the loop, while the field *outside* is *weakened*, since on that side the two fields are in opposite directions. Owing to the lateral pressure of the tubes of induction, a mechanical force acts on every part of the circuit, tending to increase its area. The movable piece of wire will be pushed towards the right.

An electric *current* in a magnetic *field* produces a *motion* in such a direction that these three quantities form a right-handed coördinate system, if taken in the above order. If the thumb,

index and middle finger of the *left* hand be held at right angles to each other, the thumb points in the direction of the field, the index finger in the direction of the current and the middle finger in the direction of the motion. Compare this "motor rule" with the "generator rule" (Art. 344).

355. Quantitative Relations for Motor. Consider a short length l of a current and a magnetic pole of strength m at a perpendicular distance d from it.

The force acting on the pole is

$$F = H'm$$

where H' is the intensity of the magnetic field due to the current. According to Laplace's law (Art. 257),

$$H' = \frac{Il}{d^2} \tag{401}$$

since in this case α is 90 degrees, and

$$F = \frac{Il}{d^2} m \text{ dynes} \tag{402}$$

where I is expressed in C. G. S. units. But according to the third law of motion, an equal and opposite force acts upon the current element. This force we may consider as due to the action between the magnetic fields due to the pole m and to the current I. The field H due to the pole at a distance d is, by equation (300), (Art. 238),

$$H = \frac{1}{\mu} \frac{m}{d^2}$$

or

$$\frac{m}{d^2} = \mu H = B \tag{403}$$

Substituting this value in (402), we have

$$F = BIl \tag{404}$$

where B is the magnetic induction at the current element. If I be given in amperes,

$$F = \frac{BIl}{10} \text{ dynes} \tag{405}$$

This equation holds for finite lengths only in case the mag-

netic induction does not vary as we pass along the wire and when the wire is at right angles to the lines of induction. Since for all gases μ is very nearly unity, and H is numerically equal to B, the force may be measured by the product HIl, but it should be remembered that in this case we have to deal with a numerical equality only.

356. The Electric Motor. An electric motor is a machine used to transform electrical energy into mechanical energy. In general, motors do not differ from generators in the details of their construction. A single loop (Art. 347) will be acted upon by a couple.

$$\mathcal{T} = Fr = 2\frac{BIl}{10} r \sin \alpha \tag{406}$$

where F is the force in dynes, r the distance of the two wires of length l from the axis of rotation, α the angle between the plane of the coil and the plane at right angles to the lines of induction, and I the current in amperes flowing through the circuit. Figure 207 shows the distribution of the lines of induction around a loop carrying a current and placed in a magnetic field at right angles to the conductors. The loop stands perpendicular to the plane of the paper, and the current flows *down* at A and *up* at B. The field is clearly distorted by the presence of the current, and a rotation in a counterclockwise direction must result if the coil be movable. The same figure would represent the distortion of the field, if the coil were used as a generator and rotated in a clockwise direction through the field.

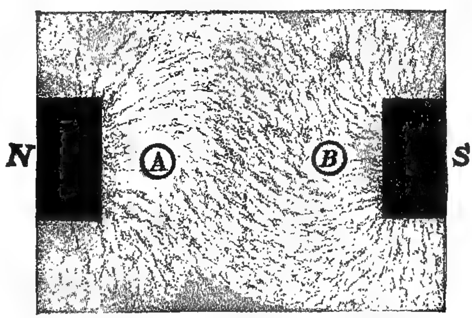


FIG. 207.

The exact correspondence between generators and motors makes it unnecessary to describe the ordinary types of motors.

The direct current motor is of the same construction as the direct current generator, and the single-phase alternating current motor is of the same form as the single-phase alternator.

Two similar single-phase dynamos may be used as generator and motor, although it is necessary first to bring the motor to the same speed as the generator. Such a motor is called a synchronous motor, but it is not self-starting, and stops when it is thrown out of step. On account of these difficulties, synchronous motors are not in general use. The induction motor (Art. 358) is now generally used in alternating current work.

357. Work done by a Motor. When the armature of a motor moves in a magnetic field, it cuts lines of induction, and therefore has induced in its coils an E. M. F., tending to *decrease* the current through the armature. This may readily be seen by applying the generator rule to the moving coil. We speak, therefore, of a *counter* E.M.F. set up in the motor. Let the difference of potential at the brushes of the motor be E, and the current through the motor be I. EI is then the rate at which electrical energy is transformed in the motor. If E' be the counter E. M. F., the energy, transformed into heat, owing to the resistance of the armature, is

$$H = (EIt - E'It) \text{ joules} \tag{407}$$

and the part transformed into mechanical energy is

$$W = E'It \text{ joules} \tag{408}$$

Disregarding the work done in overcoming friction, the mechanical power is the product of the torque $\mathscr{T}$ into the angular velocity (Art. 53), or

$$\frac{W}{t} = 2\,\pi n \mathscr{T} \tag{409}$$

where n is the number of revolutions per second. Therefore,

$$E'I = 2\,\pi n \mathscr{T} \tag{410}$$

When n is very small, for example, while the motor is being started, E' is very small, and a very large current $\frac{E - E'}{R}$ would flow through the armature, burning it out. To prevent this a re-

sistance, called a *starting box*, is placed in series with the motor, with all its resistance in the circuit before the motor is started. The resistance is cut out as the speed of rotation and hence the counter E. M. F. increase. If there be no load on the motor, that is, if the torque be small, the speed of rotation will be very high, and the counter E. M. F. allows but a small current, just sufficient to overcome the resistance due to friction, to pass through the motor. If the load be heavy, the speed decreases, and a large current flows through the machine.

***358. The Induction Motor.** If two separate coils of wire be wound upon an iron ring so that each coil covers two opposite quadrants wound in opposite directions (Fig. 208), a two-phase current sent through the two coils will produce in the iron what is known as a *rotary field*.

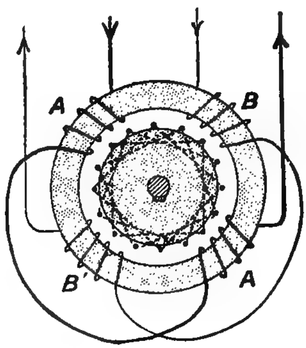


FIG. 208.

Suppose, for example, the current through A and A' to have reached its maximum in the direction indicated, while the one through B and B' is zero. The lines of induction through the iron in both halves of the ring will be directed towards B', or there will be a north pole at B' and a south pole at B. After a quarter of a period has elapsed, no current passes through A and A', but the maximum current flows through B and B' in the direction indicated. The north pole will now have advanced to A and the south pole to A'. After another quarter period, the current through A and A' has its maximum negative value, and the north and south poles have traveled to B and B' respectively. Then they advance to A' and A, and finally reach again their original position. During one complete revolution of the armature of the two-phase generator the magnetic field has made one complete revolution in the ring. Such a field is called a *rotary field*.

If a metallic disk, free to rotate about an axis in the center of the ring, be placed above the ring, it begins to whirl around

as soon as the rotary field is established. For, owing to the motion of the magnetic field, eddy currents (Art. 339) are set up in the disk, and the reaction between these and the rotary field produces mechanical forces which pull the disk around in the direction in which the magnetic field moves. *This motion is produced purely by electromagnetic induction, no electrical contact with the rotating part being needed.*

In the induction motor the *rotor* (Fig. 209) consists of two copper disks mounted on a shaft, and a large number of copper bars connecting the disks. On account of its peculiar form, such an armature is called a *squirrel-cage armature.* The induced currents flow through the bars, tending to prevent the relative motion of armature and field. In this way a mechanical torque is produced which sets the armature in rotation.

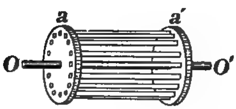


FIG. 209.

Problems

1. A cylindrical coil of 200 turns and 50 cm in length is placed with its axis parallel to the magnetic meridian in a field whose intensity is 0.19 gauss. Compute the current in amperes necessary to reduce to zero the intensity of the magnetic field at the center of the coil. *Ans.* 0.0378 ampere.

2. A steel ring having a mean radius of 8 cm and cross-sectional area of 4 cm² is wound with 60 turns of wire. When a current of 2 amperes flows in the wire, the permeability is 800. Compute (*a*) the intensity of the magnetizing field, (*b*) the induction and (*c*) the magnetic flux through the ring with 2 amperes in the wire.

Ans. (*a*) 3 gausses; (*b*) 2400 lines per cm²; (*c*) 9600 lines.

3. Determine from Fig. 177, the permeability of soft annealed iron corresponding to field intensities of 2, 4, 6, 10 and 20 gausses respectively.

Ans. 3350, 2700, 2100, 1410, 750.

4. A half-ring electromagnet is furnished with an armature, such that core and armature form a complete ring. The average diameter of the ring is 8 cm, its cross-sectional area 5 cm² and the number of turns of wire 140. If a current of 1 ampere flow through the wire, compute the magnetic flux (*a*) when the armature is pressed against the electromagnet, (*b*) when an air gap of 0.5 cm length is left between each arm of the magnet and armature. Assume that the permeability of the iron has the constant value 200, and that there is no magnetic leakage. *Ans.* (*a*) 7000 lines.
(*b*) 782 lines.

5. A wire 2 m long, placed horizontally east and west, is allowed to fall freely in a uniform magnetic field of horizontal intensity 0.19 gauss. Find (*a*) the value of the induced E. M. F. at the end of 3 sec.; (*b*) the average value of the E. M. F. during the first 5 sec; (*c*) the time elapsing before the induced E. M. F. shall be 0.001 volt.

Ans. (*a*) 0.0011172 volt; (*b*) 0.000931 volt; (*c*) 2.685 seconds.

6. A circular coil of wire, 30 cm in diameter, containing 200 turns and of 4 ohms resistance, having its ends joined together, is placed with its plane perpendicular to the earth's field of intensity 0.6 gauss. If the coil be rotated through 180° in 0.1 sec, what average current, in amperes, will be produced, and what quantity of electricity passes through the coil during this rotation? *Ans.* (*a*) 0.00424 ampere.
(*b*) 0.000424 coulomb.

7. If a bar magnet be dropped vertically through a loop of wire, it induces currents in this wire. Describe the directions of the currents.

8. The primary of a certain induction coil has 200 turns of wire, and the secondary has 20,000 turns. If the current in the primary decrease from 5 amperes to zero in 0.001 sec, compute the E. M. F. induced in the secondary, the core being of iron 10 cm long, 4 cm^2 in cross section, and of constant permeability 200. Apply formula for a very long solenoid.

Ans. 20,106 volts.

9. If a secondary of 20 turns be wound about the primary of the ring of problem 2, and the secondary have a resistance of 0.5 ohm, what quantity of electricity will pass through the secondary upon increasing the current through the primary from zero to 2 amperes, assuming the permeability to be constant? *Ans.* 0.00384 coulomb.

10. Compute the average E. M. F. in the secondary of problem 9 (*a*) with the steel inside the primary; (*b*) with air inside the primary, the current rising from zero to its maximum value in 0.05 sec. in each case.

Ans. (*a*) 0.0384 volt.
(*b*) 0.000048 volt.

11. Compute the self-inductance of a length L of a very long solenoid of cross-section A, having n turns of wire per centimeter length.

Ans. $4\,\pi\mu n^2 AL$ C. G. S. units.

12. Compute the self-inductance of a ring-shaped helix of average radius of 5 cm and cross section 4 cm^2, consisting of 500 turns of wire. Give the answer in C. G. S. units as well as in henrys. Apply formula for L, obtained in problem 11. *Ans.* (*a*) 400,000 C. G. S. units.
(*b*) 0.0004 henry.

13. Compute the energy stored in the medium inside the helix of the last problem, if it carry a current of 5 amperes: (*a*) if the medium be air; (*b*) if the medium be iron of permeability 300.

Ans. (*a*) 50,000 ergs.
(*b*) 1.5 joules.

14. If a secondary of 100 turns be wound around the helix of the last problem, compute the mutual inductance of the two coils.

Ans. (*a*) 0.00008 henry.
(*b*) 0.024 henry.

15. A rectangular loop, 20×50 cm, is rotated uniformly around its longer axis, making 400 revolutions per second. The axis is placed at right angles to a field, having an induction of 10,000 lines per square centimeter. Plot the induced E. M. F. as a function of the time.

Ans. Maximum E. M. F. is 251.33 volts.

16. A Faraday disk of radius 15 cm rotates 2400 times per minute in a field of average flux of 2000 lines per square centimeter normal to its plane. Compute the E. M. F. induced in the machine. *Ans.* 0.5655 volt.

17. If the disk of problem 16 be closed through an external circuit, what power must be applied to keep the disk in rotation, disregarding friction, and taking the total resistance equal to 5 ohms? *Ans.* 0.06396 watt.

18. A pair of wires, each having a resistance of 1.5 ohms, is used for transmitting 25 amperes with an applied E. M. F. of 2200 volts. Compute the efficiency of transmission and the drop in potential in the lines. What would be the efficiency of transmission of the same power, if the applied voltage were 550 volts? *Ans.* (*a*) 96.6 per cent; (*b*) 75 volts; (*c*) 45.45 per cent.

19. A four-pole direct current generator has 200 turns of wire on its armature. The flux from each pole is 1,250,000 lines and the speed 1200 revolutions per minute. Find the average E. M. F. induced in each turn, and the voltage developed by the machine, if half of the conductors are in series on each side of the commutator. *Ans.* (*a*) 2 volts.
(*b*) 200 volts.

20. A shunt motor has an armature resistance of 0.02 ohm. The field resistance of 55 ohms. When running on full load, the motor takes 62 amperes at 110 volts. Compute the efficiency of the motor if, in addition to the heat loss, other losses amount to 400 watts. *Ans.* 89.8 per cent.

ELECTROSTATICS

CHAPTER XLIII

FUNDAMENTAL PHENOMENA

359. Electrification. Our knowledge of current electricity dates back but a little over one hundred years to Volta's discovery of the electric cell in 1800. However, electric phenomena of a somewhat different type have been known for many centuries. Thus the Greeks knew that amber, after being rubbed, would attract light bodies, but no distinction seems to have been made between this attraction and magnetic attraction until the middle of the sixteenth century, when Cardano (1501–1576) pointed out that "*amber draws anything that is light, the magnet iron only.*" Bodies which show the same properties as amber, after being rubbed, are called *electrified bodies.*

Thus a glass rod, after being rubbed with silk and held above little bits of paper or pith balls, attracts them vigorously. They touch the rod, fly back to the table, are attracted again, and so forth.

If an electrified glass rod be brought near a pith ball, suspended by a fine silk cord, the ball will be drawn towards the glass, will adhere to it for a moment, and will then be strongly repelled. It has become electrified by contact with the glass, and the result is a *repulsion* between the two.

Experiments similar to those with the glass may be performed with a rod of hard rubber rubbed with flannel. However, it will be found that a pith ball which has been electrified by contact with the glass is not repelled by the hard rubber, but is attracted to it.

Further, it will be noticed that an electrified rod suspended so as to move freely in any direction will not assume a definite direction as a magnet will. We may therefore define *electrified bodies* as those bodies which have two characteristic properties: (*a*) they attract and repel each other with a force which is due neither to gravitation nor to mechanical action; (*b*) they show no definite orientation with respect to the geographic meridian.

360. Two Kinds of Electricity. Two-fluid Theory. If a glass rod, one end of which has been electrified by rubbing it with silk, be suspended, it will be found that the electrified end of the glass is attracted by an electrified rod of hard rubber, but is repelled by a similar rod of glass electrified by silk. An electrified rod of rubber is repelled by electrified rubber but attracted by electrified glass. This attraction or repulsion between electrified bodies was for many years explained by the assumption that, during the process of rubbing, an imponderable substance, to which the name *electricity* was given, was communicated to these bodies, and that this substance, on account of its power of action at a distance, was the cause of the mechanical force, observed between electrified bodies.

The early theory of electricity was thus very similar to that of magnetism. The different behavior of glass and hard rubber, mentioned in the last article, was generally explained by the existence of two kinds of electricity, which were designated as *positive* and *negative*. It has been agreed to call the electricity found on a glass rod, when rubbed with silk, *positive electricity*, and that on hard rubber, when rubbed with flannel, *negative electricity*.

Electricities of like kind repel, those of unlike kind attract. This fundamental law was discovered by Du Fay [1] in 1733.

We shall see later that the interpretation of the nature of electricity, as given above, must be considerably modified, but it permits of an easy and simple description of the fundamental phenomena, and lends itself readily to the solution of elementary problems.

[1] Du Fay, *Mem. de l'Acad.*, 1733.

361. Conductors and Dielectrics. We distinguished (Art. 279) between electrical conductors and insulators or dielectrics, according to the ease with which an electric current would pass through them. Experiments in electrostatics lead to the same classification. If a rod of glass or ebonite, or any dielectric, be rubbed at one end, it shows electrification only at the place rubbed; but a rod of metal, held by an insulating handle and rubbed at one end only, becomes electrified over its whole surface.

A metal rod, held by the hand and rubbed, does not show electrification, because the electricity produced escapes through the body to the earth. It follows that a conductor can retain an electric charge only if it be insulated from the earth by such substances as dry glass, ebonite, fused quartz, silk, etc.

When an insulated, neutral conductor is brought into contact with a charged conductor, some of the electricity will pass over to it, and it becomes *charged by conduction.*

362. Coulomb's Law. If two small spherical conductors be charged, they exert a definite mechanical force upon each other, which may be measured in terms of any unit of force. While investigating the attraction or repulsion between such spheres, Coulomb[1] found the law that the *force between two charged spheres is inversely proportional to the square of the distance between the centers of the spheres and directly proportional to the product of the charges*, measured in some arbitrary unit.

The similarity between this law and the law of attraction between magnetic poles (Art. 232) is so striking that it suggests immediately the existence of an influence of the medium between the charged bodies. In fact, it has been shown that such an influence exists also in this case, and that Coulomb's law in its complete form must be written

$$F = \pm \frac{1}{c}\frac{q_1 q_2}{d^2} \tag{411}$$

where q_1 and q_2 are the charges, d the distance between the points at which the charges may be considered as concentrated, and where c is a constant, characteristic of the medium.

[1] Coulomb, *Mem. de l'Acad.*, 1785, p. 569.

363. Dielectric Constant. The constant c of equation (411) is called the *dielectric constant* of the medium. It is *the measure of that property of the medium which modifies the mutual action of electrified bodies, immersed in, or separated by it.* This constant is taken as unity for a vacuum. For gases the dielectric constant is very nearly unity, for paraffine it is about 2, for mica 6, for flint glass 8. The largest value known for any substance is that of water, about 80.

The influence of the dielectric upon the capacity of a condenser has been mentioned above (Art. 316).

364. Unit Charge, Surface Density. Coulomb's law lends itself immediately to the selection of a unit of electric charge. This unit is generally defined as *that quantity of electricity which repels an equal quantity of like sign at a distance of one centimeter in vacuo with a force of one dyne.*

This charge is called the *electrostatic unit* of quantity of electricity. The student is warned not to confuse this unit with the one defined under current electricity. We deal here with static phenomena, and their relation to those of current phenomena will be considered later.

From the definition it is evident that we consider the charges as point charges, that is, as being concentrated at a point. In reality the charges are always distributed over finite surfaces, and it is often more convenient to use *electric surface densities* instead of discrete point charges. *Surface density is the charge per unit surface*, and the average surface density σ of a charge q, distributed over an area a, is given by the equation

$$\sigma = \frac{q}{a} \tag{412}$$

Unit surface density is one electrostatic unit of quantity per square centimeter.

365. The Electroscope. An electroscope is an instrument used for the detection of electric charges. In its simplest form the gold-leaf electroscope (Fig. 210) consists of a well-insulated metal rod, carrying at its lower end two thin leaves of gold foil a, b, and inclosed in a glass vessel.

If an electrified body be brought into contact with the upper end of the electroscope, the latter becomes electrified by conduction, the charge distributes itself over the metallic parts, and the leaves, being charged with electricity of the same sign, repel each other and diverge.

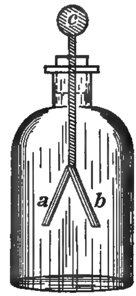


FIG. 210.

It may frequently be desired to examine the electrification of a body at some distance from the electroscope. In this case the so-called *proof plane* may be used. This consists of a small metallic disk attached to the end of a glass or rubber rod. The disk, when brought into contact with the charged body, receives a small charge which in its turn may be examined by means of the electroscope. Thus, after an electroscope has been charged positively by touching it with an electrified glass rod, the gold leaves will diverge more if the charge upon the proof plane happens to be positive, or they will collapse if the charge on the proof plane be negative. This last effect shows that positive and negative electricity, when brought upon the same conductor, tend to neutralize each other.

366. Electrification by Induction. If an uncharged, insulated conductor *B* (Fig. 211) be brought into the neighborhood of a charged conductor *A*, it will be found that the side towards *A* is charged with electricity of the opposite sign from that of *A*, and the other side with electricity of the same sign, while the central portion of *B* is not electrified, or is in a neutral condition. The existence of these charges can be shown, and their sign can be determined, by the use of electroscope and proof plane. The similarity between this experiment and the one showing induced magnetism (Art. 246) is very striking. The next experiment, however, has no counterpart in magnetism.

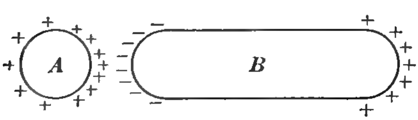


FIG. 211.

If the conductor *B*, while it is still under the influence of *A*, be touched for an instant by the hand, the charge having the

same sign as that of A escapes through our body to the earth, while the charge of the opposite sign increases in amount. If now we remove B from the influence of A, the charge upon it distributes itself over the whole surface, and B is found to be charged with electricity of only one kind. This process is called *electrification by induction.*

Let the charge thus produced be removed. The process may be repeated as often as desired, and we may produce an apparently unlimited amount of electric charge in a given body *without diminishing the original inducing charge on A.*

In magnetism we cannot produce a single pole by induction, because *no conductors of magnetism exist.*

In order to explain induced electrification, it was formerly assumed that every conductor contains an unlimited amount of both kinds of electricity which become separated under the inducing action of another charge, that of the opposite sign being attracted, and that of the same sign being repelled. When the body is connected by means of a conductor to the earth, the repelled charge will escape through the conductor, while an additional amount of charge of the opposite sign is drawn in and is then held *bound* by the attractive force of the inducing charge. This bound charge becomes free again when the inducing charge is withdrawn. A more adequate explanation will be given later.

367. Electroscope charged by Induction. When a charged body is brought towards an electroscope, the leaves begin to diverge before contact is made, owing to the charges produced by induction. If the instrument, still under the influence of the charged body, be touched by the hand, the leaves collapse, showing that their charge, which is of the *same sign* as that of the inducing body, has escaped to the earth. If now the connection with the earth be broken and the inducing charge be removed, the leaves spread apart again, showing the presence of a free charge. This charge can be shown to be of *opposite sign* to that of the inducing body.

Let us suppose the electroscope to be charged positively. Then the approach of a positive charge will produce a further

spreading of the leaves, while the approach of a negative charge, drawing the positive charge of the electroscope to the top, causes the leaves to collapse. The inductive influence upon further approach may even become so large that the leaves, after collapsing, begin to diverge again, being charged now with electricity of the same sign as that of the inducing charge.

368. Electrification of a Hollow Conductor. Let an electroscope E (Fig. 212) be connected by a fine wire to a hollow, metallic vessel V. If a positively charged body Q be introduced into this vessel, there appears upon the outside of the vessel and upon the electroscope a positive charge, and the leaves spread apart. This divergence does not change in the slightest if the charged body be moved about in the vessel. *The electrification outside is independent of the position of the charge inside.* Even if the charged body touch the wall of the vessel, no change is produced. If the body introduced into the vessel be a charged conductor, it loses all charge by touching the wall, but the external charge remains undisturbed.

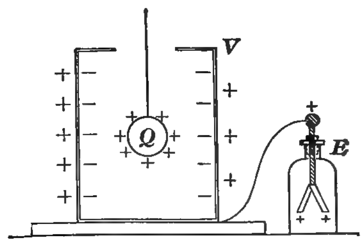


FIG. 212.

If an electroscope be placed inside a hollow conductor, as a cage made of wire netting with fine meshes, and electrically connected to it, the screen may be charged to any desired amount either from the inside or from the outside, without the slightest perceptible effect upon the electroscope. *There is no electric charge inside a charged hollow conductor ; the charge is always distributed over the outer surface.*


FIG. 213.

Faraday illustrated this fact by a striking experiment. A conical bag of netting, attached by its base to an insulated vertical ring (Fig. 213), could be turned inside out by means of a string fastened to the apex of the bag and extending axially

through the ring. On charging the bag he could show by means of a proof plane that the charge existed only upon the outside. The bag was then turned inside out and tested repeatedly, and in every case the charge was found upon the outside, although the sides of the bag had each time exchanged places.

These experiments are illustrations of the following fundamental experimental laws of electrostatics :

1. *Electric charges and electric fields within and without a hollow closed conductor are entirely independent of one another.*

2. *An electric field does not exist inside a hollow closed conductor unless there be insulated charges inside.*

369. Positive and Negative Charges always developed in Equal Amounts. Electrify two bodies by friction, as by twisting an ebonite rod *r* inside a flannel cap *c* (Fig. 214). Remove the cap by means of the silk string attached to it, and insert first the cap and then the rod into a hollow conductor joined to an electroscope (Fig. 212), being careful that no loss by conduction occurs from either. Each body, when introduced, produces the same divergence of the leaves, but the charges are of opposite sign. If both bodies be introduced at the same time, no effect can be observed.

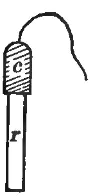


FIG. 214.

Again, after an insulated charge has been placed in the vessel, the leaves of the electroscope show a certain divergence. If now another *insulated*, but *uncharged*, conductor be introduced, it will be electrified by induction, but the divergence of the gold leaves remains the same as before.

Both experiments show that in *electrification by friction, as well as by induction, positive and negative charges are always developed in equal amounts.* It may be shown that this is always the case, regardless of the manner in which the electrification is produced.

370. Distribution of an Electric Charge. If a metallic chain be placed upon a conducting plate which is connected to an electroscope, it will be seen that when the plate is charged, the electroscope shares the charge of the system. If now one end

of the chain be raised from the plate by means of a silk thread, the leaves of the electroscope tend to collapse, thus showing that *we have decreased the surface density of the charge by increasing the effective surface of the conductor.* When the chain is returned to the plate, the leaves resume their original position, thus indicating that the original surface density of the charge has been restored.

Again, *the surface density of a charge upon a conductor of irregular outline varies with the curvature of the surface of the conductor.* Thus, in an egg-shaped insulated conductor, we shall find, upon testing the charge by means of a proof plane, that the plane removes the greatest charge from the conductor when applied at the smaller end, or since the area of the plane is constant, we find that the density of the charge is greatest at the point of maximum curvature. As the small end of the conductor becomes more and more sharp, the density of the charge increases, until finally the charge escapes into the surrounding air, by means of the convection of the individual, highly charged air particles. This fact explains the effect of points in discharging insulated charged conductors. For this reason all sharp points or edges are to be avoided in electrical machines or other apparatus intended to retain a charge of electricity. A coat of dust is equally efficient in discharging a charged body.

371. The Frictional Machine. This type of machine was formerly used for the continuous production of electric charges. It consists of a large glass disk, the various parts of which during rotation come into close rubbing contact with a pair of leather cushions, covered with a tin amalgam. Upon rotating the disk, both surfaces of the plate become positively electrified by friction, and the electricity thus produced is collected upon an insulated conductor by means of a system of conductors armed with metallic combs whose sharp points lie near the surface of the plate on either side.

Since the invention of the much more powerful influence machines (Art. 373), the frictional machines have gone entirely out of use, but are still found in collections of old apparatus.

372. The Electrophorus. The electrophorus, an instrument invented by Volta, is the simplest apparatus for producing electricity by induction. It consists of a plate or flat cake *A* (Fig. 215) of hard rubber, rosin or some other dielectric, and of a metallic disk *B*, to which is attached an insulating handle *H*. The dielectric plate is usually placed upon a metallic plate *S*, or cast in a shallow metal dish.

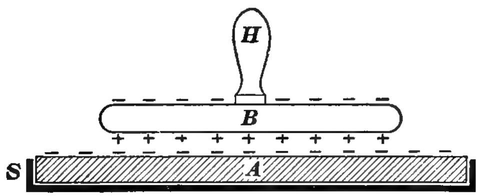


FIG. 215.

After the dielectric has been excited by whipping it with a piece of catskin, it is negatively electrified, and induces a positive charge upon the under side and a negative charge upon the upper side of the metallic plate when placed upon it. The irregularities of the surface of the lower plate prevent loss of electricity by conduction. If now the conductor be touched at any point by the finger, the induced free negative charge flows off to the earth through the body of the operator, while the positive remains bound on the plate. After the finger is removed, the metallic plate may be lifted off by the handle, and the electricity upon it, becoming free, is available for charging other bodies.

***373. Influence Machines.** The first successful influence machine, capable of producing relatively large charges, was described by Toepler[1] in 1865. These machines have since been greatly improved by Holtz, Wimshurst and others. Large amounts of electricity may readily be produced by means of these machines. Such machines are usually employed for performing a series of interesting and striking experiments, illustrating the principles and laws discussed in the previous paragraphs.

The action of the influence machine is based upon the electrification of its rotating parts by induction. The action is somewhat complicated, and different in different types. The student

[1] A. Toepler, *Pogg. Ann.* 125, p. 469, 1865.

who wishes to familiarize himself with any one of these machines will find description and explanation of its action in several of the larger handbooks on Physics.

374. The Electric Spark. If the knuckle be brought close to the charged plate of an electrophorus, a small spark passes across the gap. The terminals of an influence machine become highly charged with electricities of opposite sign, and sparks of much greater length and volume may be obtained in rapid succession from one of these machines. When received upon the body, these sparks are painful and sometimes dangerous.

The sparks are caused by a recombination of positive and negative electricity, and become visible to the eye by the heating to incandescence of the medium through which the spark passes. At the same time, owing to the sudden expansion of the heated gas and to its subsequent contraction, the sound characteristic of the spark is heard.

375. Spark and Electric Current. We have already (Art. 333) identified the spark produced in an inductive circuit with an electric current. It is easy to show that the spark, due to the combination of positive and negative electricity, is of the nature of a current, since by passing it through a helix surrounding a steel needle, the needle will be found to be magnetized, as if the helix had been traversed by an electric current.

The following experiment will remind the student of one described in Art. 282. Place in series with the terminals of an influence machine two fine platinum wires touching upon the opposite ends of a strip of filter paper, moistened with a solution of sodium sulphate, and colored with litmus or extract of purple cabbage (Fig. 216). The spark gap may be placed anywhere in the circuit. After the passage of a few sparks, the side A of the paper, connected with the positive end of the machine, will have become red; that is, the wire through which the positive electricity entered the solution is the anode.

FIG. 216.

The effect is, therefore, the same as if an electric current had

passed through this modified electrolytic cell in the same direction in which the positive electricity is assumed to travel across the strip of paper.

From these experiments we may conclude that a spark, passing across a gap in a conducting circuit, may be considered as an indication of an electric current flowing through the circuit in the same direction as that in which the positive charge passes across the gap. Since an electric current is the time rate of transfer of electricity, and since in this case the quantity as well as the time of transfer is very small, we may write

$$I = \frac{dq}{dt} \tag{413}$$

376. Lightning and Lightning Rods. The similarity between the form of the sparks produced by electrostatic machines and the lightning flash suggested to the early workers in electricity that lightning is but a mighty electric spark. This conclusion was confirmed by Franklin in his famous experiment, 1752, in which he drew electric sparks from passing clouds by means of a kite attached to a wet string.

Lightning flashes are thus seen to be discharges between oppositely charged conductors. They may occur between two clouds or between a cloud and the earth.

In order to protect houses from lightning, lightning rods are frequently used. They are simply conductors of large surface area, leading from the top of the house to the earth and connected to a large sheet of metal, buried in moist earth. All metallic masses in the house should be metallically connected to this conductor. At its upper end the lightning rod is provided with one or more sharp points. These act to a certain degree as equalizers by allowing electricity of the opposite sign from that of the cloud to escape into the air, and thus to neutralize in part the charge of the cloud. The lightning rod will afford no protection unless all electrical connections are in perfect condition.

***377. The One-fluid Theory.** In the preceding article the fundamental facts of electrostatics have been presented from the

point of view of the two-fluid theory. The difference between positively and negatively charged bodies has been explained by assuming the existence of two distinct, unlike kinds of charges. But it should be noted that this is not the only possible interpretation.

Thus Benjamin Franklin (1706–1790) proposed a one-fluid theory, which assumes that only one kind of "*electric fire*" exists. According to his view, a body is positively charged if it possess more than its normal share of this fluid, and negatively charged if it possess less than its normal share. In fact, we owe the terms *positive* and *negative* to Franklin. From this point of view, an electric spark or current must be considered, not as a combination of two unlike charges, but as a transfer of electricity in one direction only. It may be compared to the flow of a gas from a vessel under high pressure to one partly evacuated.

Recent experiments by Nipher point strongly toward a one-fluid theory. The electron theory also assumes that an electric current is mainly, if not entirely, due to a transfer of *negative* charges through the conductor. We shall return to this theory in a subsequent chapter.

CHAPTER XLIV

THE ELECTROSTATIC FIELD

378. Electrical Theories. It has been pointed out that some of the simplest electrostatic phenomena may be easily explained by the assumption of action at a distance between two hypothetical electric fluids, called positive and negative charges. But, as in magnetism, scientists were forced to give up the *action-at-a-distance theory* and accept as a working hypothesis the *ether-strain theory*, as first introduced by Faraday and later developed by Maxwell (1831–1879). Maxwell showed that the ether-strain theory, when applied to disturbances in dielectrics, leads to the conclusion that light is nothing but an electromagnètic disturbance in the ether, and he thus founded the electromagnetic theory of light.

In recent years many new experimental facts have been discovered, especially in the field of electrical discharges in gases and in radioactivity, which have shown Maxwell's original theory to be inadequate, and which have led to a modification of this theory, generally known as the *electron theory*. This is to a certain extent a combination of the older theory with Maxwell's theory, and, while not complete in all its details, it bids fair to become the leading theory. A subsequent chapter will be devoted to its study. The ether-strain theory suffices, however, for a satisfactory explanation of all electrostatic phenomena.

379. The Ether-strain Theory. The fundamental concepts of this theory have been fully given in a preceding chapter under Magnetism. According to Faraday and Maxwell, electrostatic phenomena are due to a strained condition of the medium between electrically charged bodies. To make this

idea more concrete, Faraday introduced the concept of strain tubes, which are supposed to extend *from* a positive charge *to* a negative charge. Thus we may think of a large number of such tubes as originating between a glass rod and the silk, when the two are rubbed together, and to be stretched and spread out into space when the two are separated from each other.

An insulated conducting sphere, charged with positive electricity, is thus surrounded by strain tubes, starting out in all directions from the surface of the conductor. The tubes have a tendency to *shorten and to exert a lateral pressure upon each other.*

By the same reasoning as that given in (Art. 244) the fundamental law of attraction and repulsion between electrified bodies (Art. 360) is derived.

380. Conductors in an Electrostatic Field. One fundamental difference between magnetic and electrostatic phenomena has already been mentioned (Art. 366), namely, that no *magnetic conductors* are known. An electrical conductor is unable to support an electric strain. It allows, so to speak, the ends of the strain tubes to slip along its surface. It is, therefore, impossible, in an electrostatic field, for tubes to start and end upon the same conductor.

Thus let us suppose that at a given time, before equilibrium has been established, some tubes do extend from some points on the inside of a hollow conductor to other points. The tubes will contract and finally disappear, and thus leave the interior entirely free from strain tubes. In other words, there can be no electrostatic field inside a hollow conductor (Art. 368).

Again, if a conductor be brought into an electrostatic field, it cuts the tubes asunder, and the new ends of the tubes slip along the surface, until equilibrium is established between the lateral pressure between neighboring tubes and the tendency of each tube to shorten. The phenomena of induced electrification are thus easily explained. Figure 217 gives the complete representation of this state of the medium, which was entirely disregarded in the corresponding figure (Fig. 211). Thus (Fig. 217)

we see that to every tube terminating upon a negative charge on the left-hand side of B there corresponds a tube starting out from a positive charge on the right-hand side of B. The equality of positive and negative charges, produced by induction (Art. 369), is a direct consequence of this theory.

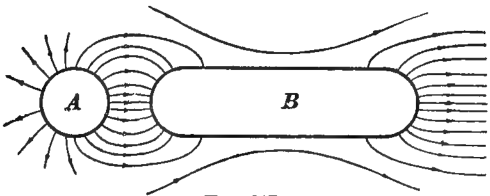


FIG. 217.

381. Further Applications. The conditions existing in and about an electrophorus after the metal plate has been placed upon the dielectric are represented in Fig. 218. It is clear that when the upper plate is touched by the finger, all tubes between the plate B and the earth disappear through the body, and there remain only the tubes between the plate and the dielectric, or, the plate is charged positively.

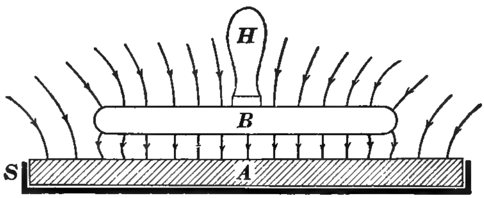


FIG. 218.

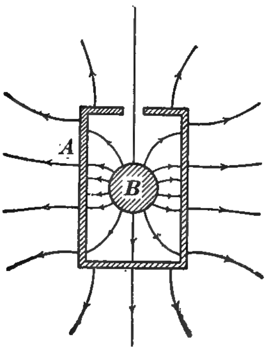


FIG. 219 *a*.

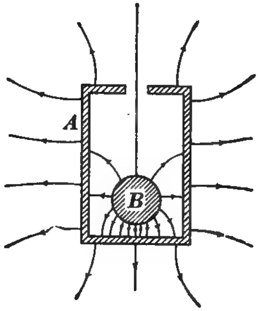


FIG. 219 *b*.

Again, when a charge is brought into a hollow conductor, the tubes inside the conductor are completely separated from

those on the outside. The two fields are entirely independent of each other, as illustrated (Figs. 219 *a*, 219 *b*). The distribution of the tubes outside the vessel does not depend upon that on the inside. If the outside be connected to the earth by a conductor, the tubes outside the vessel, that is, the field, will disappear. If the conductor B touch the wall of the vessel, it is discharged, the inside field disappears, while the field outside remains undisturbed.

From these applications of the theory to the various phenomena described in the preceding chapter, it it seen that *electric charges are the manifestations of a discontinuity of electric strain tubes.*

382. Intensity of Electric Field. When a charge q is brought into an electric field, a force F acts upon this charge, no matter to what causes the field is due. The intensity of the field E is related to the force by an equation analogous to equation (297),

$$F = Eq \tag{414}$$

The intensity of an electric field at a given point may therefore be measured by the force acting upon a charge brought to that point, provided that the presence of the charge does not appreciably disturb the whole field.

Unit intensity of field at a point is that field intensity which produces a force of one dyne when unit charge is brought to that point.

If the field at a point P be due to a charge q' at a distance d from P, the last equation, combined with Coulomb's law,

$$F = \pm \frac{1}{c} \frac{qq'}{d^2}$$

leads to the relation

$$E = \pm \frac{1}{c} \frac{q'}{d^2} \tag{415}$$

The intensity of an electric field at any point is a vector quantity, and has the same direction and sense as the force acting upon a positive charge brought to that point.

The intensity of an electric field is by many writers called

the electric force, but, as is easily seen, it is not a mechanical force, since *no such force exists at the point in question, unless a charge be placed there.*

383. Electric Induction. The strained condition of the medium between electrified bodies has been represented, as in magnetism, by tubes or lines of induction. In a vacuum, where c equals unity, the *intensity of the field*, and the *electric induction* D, where

$$D = cE \tag{416}$$

are both *unity* at a distance of one centimeter from a *unit point charge.* They are both q units at a distance of one centimeter from a point charge q. If we agree to represent unit induction by one line per square centimeter, *the total number of lines of induction passing through a sphere of unit radius, described about a point charge of* q *units as center, is* $4\pi q$. The number of lines of induction is independent of the medium surrounding the charge, while the intensity of the field is inversely proportional to the dielectric constant of the medium.

As in magnetism (Art. 242), *induction* corresponds to *a strain* in an elastic body, while the *intensity of the field* corresponds to *a stress.*

384. Work done in moving a Charge. Since a charge q, placed in an electric field, is acted upon by a force F, where

$$F = Eq$$

work must be expended in moving the charge from B to A in a direction opposite to that of the force, and work is done by the electric field when it displaces the charge from A to B in the direction of the force.

If the field be uniform, that is, if E be a constant for all points of the field, the work is evidently

$$W = Fs = Esq \tag{417}$$

where s is the distance AB in the direction of the force through which the charge has been moved.

If the field be not uniform, the path of the charge must be divided into a number of small parts ds, along each of which

the intensity may be considered as uniform. The total work done is then the sum of all such terms as $Eqds$, or

$$W = \Sigma Eqds \tag{418}$$

This summation requires the use of calculus.

385. Electrostatic Difference of Potential. Equation (418) may also be written

$$W = (V_A - V_B)q \tag{419}$$

where $V_A - V_B$ is called the *difference of potential*, or fall of potential, between the points A and B, between which the charge has been moved.

The difference of potential between the points A and B is

$$V_A - V_B = \frac{W}{q} \tag{420}$$

and is therefore *the work per unit charge required to move the charge from one point to the other.* It is numerically equal to the work if the charge be one electrostatic unit. *Two points are said to have unit difference of potential when one erg is required to move one electrostatic unit of charge from one point to the other.*

386. Electrostatic Potential. Equations (417) and (419) show that

$$V_A - V_B = Es \tag{421}$$

which may be written

$$\frac{V_A - V_B}{s} = E \tag{422}$$

$V_A - V_B$ is thus seen to be a change in value of a physical quantity, called electrostatic potential, *whose rate of variation in space equals the intensity of the field.*

Since the direction of the field has been assumed to be from A to B, V_A is larger than V_B, and $V_A - V_B$ is a positive quantity, while $V_B - V_A$ is negative. While difference of potential may be either positive or negative, according to the position of the two points with respect to the direction of the field, *potential and difference of potential are scalar quantities*, just as work and electric charge.

If the field be not uniform, equation (422) can only be applied over very short distances ds, and must be written

$$\frac{dV}{ds} = -E \tag{423}$$

where the negative sign indicates that dV is a *decrease* of poten tial, when taken in the positive direction of E.

387. Potential at a Point due to a Charge. In order to be able to assign a definite numerical value to the potential at a point in a field, it is necessary to choose arbitrarily some point as being at zero potential. At an infinite distance from any electric charge there must be zero potential. The potential at any point may then be calculated by finding the work required to bring unit positive charge from infinity up to this point, the dielectric constant of the whole field being c.

By the use of calculus it may be shown that the potential V at a point P, due to a charge q concentrated at a distance d from P, is

$$V = \frac{1}{c}\frac{q}{d} \tag{424}$$

in which the sign of q determines whether V is positive or negative.

The potential at a point outside a sphere, charged uniformly with a quantity q, is the same as if the total charge of the sphere were concentrated at its center, and the above equation holds in this case also, where d denotes the distance of the point in question from the center of the sphere.

For all practical purposes the potential of the earth is assumed to be zero, just as the sea level is chosen as the standard level from which to measure height.

388. Superposition of Electric Fields. If a number of concentrated charges and their positions be given, the condition of the field at any point in space, characterized by the field intensity and potential at that point, is determined by finding the sum of the fields produced by the individual charges.

The field intensity, due to one of these charges q_n, is, by equation (415),

$$E_n = \frac{1}{c}\frac{q_n}{d^2_n}$$

To obtain the total field intensity at the point in question, we must form the *vector sum* (Art. 12) of the component fields.

The potential, due to one of these charges q_n, is, by equation (424),

$$V_n = \frac{1}{c}\frac{q_n}{d_n}$$

The potential, due to all the charges, is simply the *algebraic sum* of the potentials due to the individual charges.

389. All Points of a Conductor in an Electrostatic Field at the Same Potential. If two points of the surface of a conductor were at different potentials, there would be a definite field intensity between the points according to equation (422), and strain tubes would extend from one point to the other. But (Art. 380) a conductor in an electrostatic field is unable to support an electric strain, and the strain tubes contract indefinitely. Equilibrium can exist only when all points of the conductor are at the same potential.

While a conductor under the inductive action of a charge may be electrified positively at one end and negatively at the other, the potential over its whole surface must necessarily be the same, after electrostatic equilibrium has been reached.

390. Potential of a Spherical Conductor due to its Own Charge. The charge on a conducting sphere, at a great distance from any other charge, is uniformly distributed over the surface of the sphere. Since there is no field inside the conductor (Art. 368), there is no difference of potential inside the sphere, and the center of the sphere must be at the same potential as the surface. In order to find the potential of the sphere, it is therefore sufficient to find the potential at its center. The dielectric constant of the space assumed to be c, the potential at the center due to any small portion dq of the charge on the surface is

$$dV = \frac{1}{c}\frac{dq}{r} \tag{425}$$

where r is the radius of the sphere. The potential, due to the whole charge, is the sum of the potentials produced by each portion of the charge on the sphere, or,

$$V = \frac{1}{c}\frac{\Sigma dq}{r} = \frac{1}{c}\frac{q}{r} \tag{426}$$

391. Equipotential Surfaces. An equipotential surface is a surface on which the potential is the same for every point. It is easily seen that the equipotential surfaces around an isolated charged sphere are spherical surfaces.

For example, let a conducting sphere of 0.5 cm radius be charged with 4 electrostatic units of electricity, and let the surrounding medium be air, *i.e.*, let c be unity. The potential of the sphere is then 8 units. The points whose potential is 7 are all located on a sphere whose radius is $\frac{4}{7}$ cm, and the following equipotential surfaces, each of a potential smaller by one unit than the preceding, have radii equal to $\frac{4}{6}$, $\frac{4}{5}$, 1, $\frac{4}{3}$, 2, 4 cm and infinity (Fig. 220).

Since the intensity of the field at any point is (422),

$$E = \frac{V_A - V_B}{s}$$

and the difference of potential in the successive surfaces in Fig. 220 is constant and equal to unity, it is evident that the

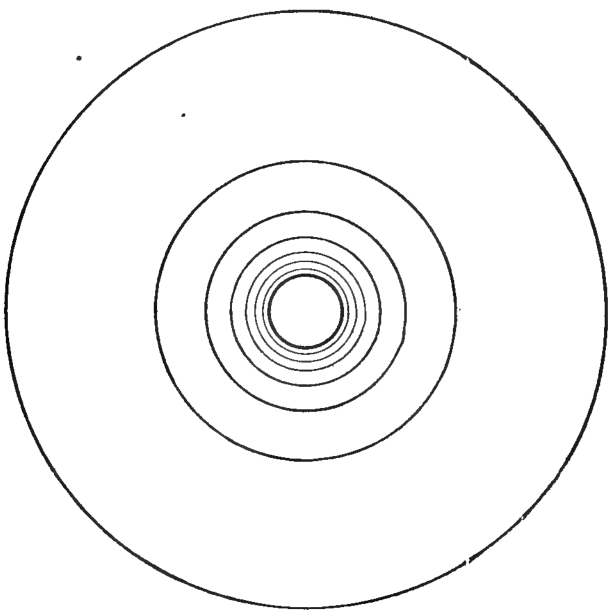


Fig. 220.

field intensity is inversely proportional to the distance between two neighboring equipotential surfaces.

The method of representing equipotential surfaces is very similar to that employed in maps to represent difference in level. The closer together the contour lines are at any point, the steeper is the slope.

392. Variation of an Electrostatic Field and Current Electricity. Whenever conductors at different potentials are connected by a wire, the strain disappears, and the electrical energy stored in the medium is transformed into other forms of energy, usually into heat.

The terminals of an electric cell, when on open circuit, are at different potentials, and we may imagine strain tubes to extend from one terminal to the other. If the terminals were simply two insulated metallic plates, the strain tubes and electric energy would disappear on connecting them by means of a wire. An extremely small quantity of electricity would flow through the circuit, accompanied by the development of a small amount of heat. But the chemical reaction going on in the cell furnishes a constant supply of energy which keeps the terminals of the cell at a definite difference of potential, and thus, by continually establishing new strain tubes, makes the flow of electricity a continuous process. The resulting phenomena have been studied in the chapter on current electricity.

From this point of view the phenomenon of an electric current is not restricted to the substance of the conductors carrying the current, but takes place mainly in the space about the circuit. The conductors form only that part of the system in which an amount of heat appears equivalent to the energy constantly given up by the electric field. In a dynamo the energy needed for the continuous maintenance of a difference of potential along the metallic circuit is furnished by the mechanical energy spent in driving the generator.

There must be some close connection between the ether and the material substance of a conductor. The ether-strain theory does not furnish an explanation for this. An attempt to do so is made by the electron theory, which will be treated in a subsequent chapter.

CHAPTER XLV

ELECTROSTATIC CAPACITY

393. Capacity of an Insulated Conductor. If an insulated conductor be charged, it will be found that the potential of the conductor is strictly proportional to the charge, or, in mathematical terms,

$$q = CV \tag{427}$$

where the proportionality factor C is called the *capacity of the conductor*. It may be defined as the ratio of the charge of the conductor to the potential produced by the charge. If both q and V be expressed in electrostatic units, the capacity is also given in these units.

From the above equation the capacity of an insulated sphere may easily be calculated. The potential of a charged sphere, due to its own charge, is by equation (426)

$$V = \frac{1}{c}\frac{q}{r}$$

Substituting this value in the equation for C, we have

$$C = \frac{q}{V} = cr \tag{428}$$

The capacity of a spherical conductor is therefore proportional to the radius of the sphere and to the dielectric constant of the surrounding medium. From equation (428) it appears that the capacity of an insulated spherical conductor *in air*, *is numerically equal to its radius.*

394. Potential measured by the Electroscope. We have seen (Art. 365) that an electroscope may be used to detect the presence of an electric charge upon a conductor by transferring to it a small part of the charge from the conductor. Since the potential

of the electroscope is proportional to its charge, the divergence of the gold leaves also indicates the potential of the instrument.

If now an electroscope be connected by a wire to a large charged conductor B, a small fraction of the charge on B will be transferred by conduction to the electroscope, until its potential is equal to that of B. But the potential of B will hardly be affected. In this case the divergence of the gold leaves does not measure the *charge* upon B, but the *potential* and the electroscope may therefore be used to detect any variation in the potential of the charged body with which it is connected.

395. Effect of Neighboring Conductors. If a large insulated metal plate, connected with an electroscope, be charged, the divergence of the gold leaves serves as a measure of the potential of the plate. If now a similar plate connected to the earth be brought near and parallel to the first plate, the divergence of the gold leaves becomes appreciably smaller. The effect *increases* as the distance between the plates *decreases*. This proves that the potential of the charged plate has been decreased by the presence of the second plate. Since the *charge* has not been changed, the *capacity* of the plate has evidently *been increased*. Upon removal of the second plate, the potential of the charged plate rises to its original value, owing to a decrease in the capacity of the plate.

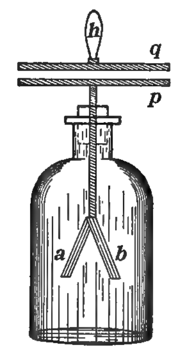


FIG. 221.

Instead of connecting the electroscope to a separate plate, it may be furnished with a large plate p, placed on the rod holding the gold leaves (Fig. 221). Upon the approach or the removal of another similar plate q, connected to the earth, the varying divergence of the gold leaves shows clearly the change in the capacity of the plate of the electroscope. Such an instrument, furnished with a system of two plates, is called a condensing electroscope.

The phenomena just described may be easily explained by the ether-strain theory. With the upper plate removed, lines

of electric induction pass from the charged electroscope in all directions to the earth, many of them from the gold leaves. This results in a divergence of the leaves, owing to the mutual repulsion of the lines (Art. 379). When the second plate, which, by virtue of its connection with the earth, is at zero potential, is brought near, the lines crowd into the space between the plates and leave only a weak field outside the plate (Fig. 222). The work required to bring a charge from the earth to the plate is now much smaller than it was before, or the potential of the charged plate has decreased.

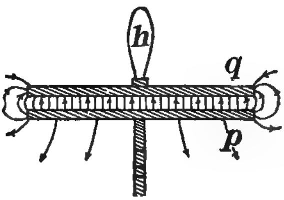


FIG. 222.

396. Capacity and Charge of a Condenser. The two plates of the last article form an electric condenser (Art. 313), whose capacity was defined in a manner very similar to that of a conductor, namely, by the equation

$$C = \frac{Q}{E} = \frac{Q}{V_1 - V_2} \tag{429}$$

where Q denotes the quantity of electricity, as defined in current electricity, and E the difference of potential between the terminals of the condenser. From the electrostatic point of view, the charge of a condenser consists of both positive and negative charges, of equal magnitude. Their sum would be zero. In order to express the capacity of a condenser in electrostatic units, it is therefore necessary to consider only the charge upon one plate, or one side of the condenser. The distinction between charges of different sign is thus abandoned, and capacity is always a positive quantity.

No fundamental difference exists between the charge of a condenser produced by an electrostatic machine and that produced by an electric cell or other current generator. This is readily proved by connecting for an instant the two plates of a condensing electroscope, separated by a thin sheet of mica, to the terminals of an electric battery. The plates will be charged

to a difference of potential very much smaller than that produced by electrostatic machines, and there will be no noticeable effect upon the gold leaves. But if now the upper plate be removed, the capacity of the instrument is greatly *decreased*, and the gold leaves diverge, owing to the *increase* of potential. The electric strain, which before was practically restricted to the space between the plates, is now distributed over the whole instrument, and the lines of induction proceeding from the gold leaves force them apart.

397. Capacity of a Spherical Condenser. A spherical condenser consists of two concentric spherical conducting shells separated by a thin dielectric. Let the outer radius of the inner conductor A (Fig. 223) be r_1, and the inner radius of the outer conductor B be r_2. If one of the conductors, for example B, be connected to the earth, and the other, A, be charged with a quantity $+q$, a quantity $-q$ is induced upon the inside of B, while the potential of A is raised to V.

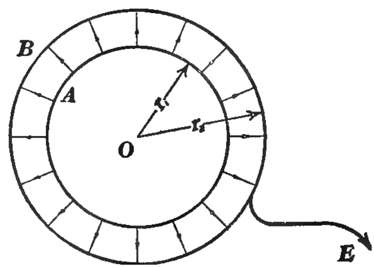


FIG. 223.

In order to evaluate the difference of potential between A and B, it is only necessary to find the potential at the common center O, since this is equal to that of A (Art. 390), and the potential of B is zero, since it is connected to the earth. The potential at O is equal to the sum of the potentials V' and V'' due to $+q$ on A and to $-q$ on B. From equation (426) we have

$$V' = +\frac{1}{c}\frac{q}{r_1}$$

and

$$V'' = -\frac{1}{c}\frac{q}{r_2}$$

Therefore

$$V_1 - V_2 = V = V' + V'' = \frac{q}{c}\left(\frac{1}{r_1} - \frac{1}{r_2}\right) = \frac{q}{c}\frac{r_2 - r_1}{r_1 r_2} \qquad (430)$$

The capacity of the spherical condenser is thus

$$C = \frac{Q}{V_1 - V_2} = \frac{q}{V} = c\,\frac{r_1 r_2}{r_2 - r_1} \tag{431}$$

If the radii r_1 and r_2 be nearly the same, equal to r, and if t be the distance between the shells, the last equation may be written

$$C = c\,\frac{4\,\pi r^2}{4\,\pi t} = c\,\frac{A}{4\,\pi t} \tag{432}$$

where A is the surface of the sphere of radius r.

398. Capacity of a Plate Condenser. The capacity of a plate condenser may be found from that of a spherical condenser. The lines of induction, and consequently the charges, are evenly distributed over the surfaces of the conductors (Fig. 223). This means that the capacity of any portion of the condenser is proportional to its surface. If now we assume that the radii of the spheres become very large, any section cut from this condenser will be a plate condenser, and consequently

$$C = c\,\frac{A}{4\,\pi t} \tag{433}$$

where A is the area of the dielectric between the plates and t its thickness.

It should be noted that, in general, some of the lines will spread out beyond the edges of the condenser (Fig. 222).

For small values of t this effect will be negligible, and equation (433) will be very nearly correct for condensers whose conducting plates are separated by very thin sheets of dielectric.

399. Leyden Jars. The condensers most frequently employed in electrostatic experiments are the so-called Leyden jars, invented by von Kleist[1] in 1745. A Leyden jar consists of a glass jar coated to a certain height inside and outside with tin foil, the remaining free part of the glass being covered with shellac (Fig. 224). A metal rod carrying a brass knob passes through the wooden cover of the jar and makes connection with the inner tin-foil coating through a fine chain.

FIG. 224.

The jar is charged by holding it in the hand, thus

[1] Von Kleist, *Abh. d. Naturf. Ges. Danzig*, 2, p. 407, 1745.

connecting the outer coating with the earth, and bringing the knob into contact with one of the terminals of an electrostatic machine. The glass between the layers of tin-foil then becomes the seat of a strong electric strain. The jar is best discharged by the discharger (Fig. 225), consisting of a jointed brass rod provided with a glass handle. One of the knobs is laid against the outer coating, and the other is brought close to the central rod of the jar. A bright spark breaks across the gap and thus relieves the electric strain.

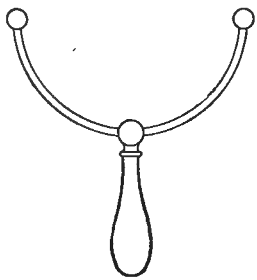


FIG. 225.

Since glass shows the phenomenon of electric absorption, the strain does not disappear entirely upon the first discharge, so that a succession of sparks, each weaker than the preceding, may be obtained from a strongly charged jar.

400. Influence of the Dielectric upon Capacity. The influence of the dielectric upon the capacity of a condenser (Art. 316) may readily be shown by means of the condensing electroscope. Let the instrument be charged while the upper plate is connected to the earth and supported a few centimeters above the lower plate. If now a dielectric sheet, as a plate of glass, sulphur or paraffine, be introduced between the plates, the divergence of the gold leaves decreases, showing that the potential has *decreased*. Since the charge has not changed, the capacity of the condenser must have *increased*, owing to the substitution of the dielectric plate for air. Upon the withdrawal of the dielectric plate, the gold leaves assume their former positions.

***401. Electrostatic Energy.** If an insulated conductor, originally without charge and at zero potential, be charged by placing upon it a series of small charges, its potential will rise in proportion to the charge, until it finally reaches the value V, when the total charge q has been placed upon it. By making the steps quite small, the process of charging may be made practically uniform. It is obvious that the average potential during

the time of charging is $\frac{1}{2}V$, and that the work W expended in charging the conductor is the same as if the total charge q had been carried from a point of zero potential to one of potential $\frac{1}{2}V$. From equation (419) we have in this case

$$W = \frac{1}{2}Vq = \frac{1}{2}CV^2 = \frac{1}{2}\frac{q^2}{C} \tag{434}$$

In a similar manner the work necessary to charge a condenser with a quantity Q to a difference of potential $V_1 - V_2$ may be considered as the work necessary to carry the charge from one plate of the condenser to the other, the average difference of potential being $\frac{1}{2}(V_1 - V_2)$. Thus we obtain

$$W = \frac{1}{2}(V_1 - V_2)\,Q = \frac{1}{2}C(V_1 - V_2)^2 = \frac{1}{2}\frac{Q^2}{C} \tag{435}$$

If the electric quantities be expressed in C. G. S. units, the work is given in ergs; if the electric quantities be expressed in coulombs, volts and farads, the work is given in joules.

Work is thus transformed into electrostatic energy stored in the dielectric of the charged condenser. Upon discharge, it generally appears as heat in the spark, or in the conductor through which the discharge takes place. In special cases, however (Art. 405), a small part of the electrostatic energy appears as energy of radiation.

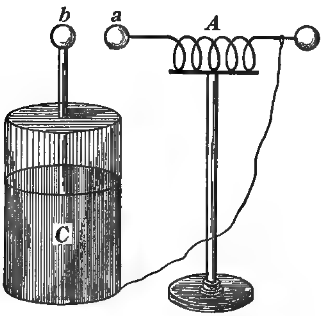


FIG. 226.

402. Oscillatory Discharge of a Condenser. In 1827 Savary[1] observed that when an electric spark from a Leyden jar passed through a helix A (Fig. 226) surrounding a needle, the needle was sometimes magnetized in one sense, sometimes in the opposite sense. He explained this by assuming the discharge to have an oscillatory character.

In 1853 Lord Kelvin[2] (1824–1907) showed from mathematical

[1] Savary, *Pogg. Ann.* 10, p. 100, 1827.

[2] Thomson (Kelvin), *Phil. Mag.* 5, p. 393, 1853.

considerations that the discharge of a condenser through a circuit, containing resistance, capacity and self-inductance, may indeed be oscillatory. The conditions under which electric oscillations in such a circuit are produced are: (*a*) that the discharge shall be very sudden, as in the case of a spark; (*b*) that the self-inductance bear a definite relation to the resistance and the capacity of the circuit, namely, that $R^2 < \frac{4L}{C}$.

In this case the *frequency* of oscillation n is

$$n = \frac{1}{2\pi}\sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}} \tag{436}$$

or, if R^2 be negligible in comparison with $\frac{4L}{C}$,

$$n = \frac{1}{2\pi}\sqrt{\frac{1}{LC}} \tag{437}$$

and the period T is

$$T = 2\pi\sqrt{LC} \tag{438}$$

where the electrical quantities are all measured in the same system of units.

If R^2 be larger than $\frac{4L}{C}$, no electrical oscillation can be obtained, and the discharge becomes aperiodic.

In 1857[1] Feddersen examined the discharge from a Leyden jar by means of a rapidly revolving mirror, and observed, instead of a single spark, a succession of flashes decreasing in brightness.

In 1862 Paalzow[2] modified the experiment of Feddersen by discharging the jar through a Geissler tube, and examining the appearance of this tube in the rotating mirror. In such a tube the discharge of a condenser produces a flash of light in the rarefied gas in the tube. This light is reddish at the anode and bluish at the cathode. Now Paalzow found that the tube, when illuminated by a discharge from the jar, showed in the rotating mirror not only a series of bright images of the tube, but that

[1] Feddersen, *Pogg. Ann.* 103, p. 69, 1858.
[2] Paalzow, *Berl. Ber.* 1862, p. 152.

these images were at either end ***alternately red and blue***, thus proving conclusively the oscillatory character of the discharge.

A mechanical analogue may be helpful for a clearer understanding of this phenomenon. Let a weight supported by an elastic spring be pulled down, producing an elastic strain in the spring. If now the weight be suddenly released, it will make a number of oscillations up and down about its position of equilibrium. In a similar manner oscillations are set up in the ether when the electric strain is suddenly released.

If the weight be placed in a viscous medium, the number of oscillations will be greatly decreased, or, if the viscosity of the medium be large, no oscillations will occur, the system returning slowly to its position of equilibrium. This last case corresponds to the aperiodic discharge of a condenser through a circuit of high resistance.

***403. The Singing Arc.** The existence of oscillations in an electric circuit may be shown by the following interesting experiment. If a circuit containing a capacity C of several microfarads (Fig. 227) and a small self-inductance L be connected in parallel with a direct current arc A between solid carbon tips, and if the current be carefully adjusted to about 3 amperes, the arc will emit a clear, musical tone whose pitch may be varied by changing either the capacity or self-inductance or both.

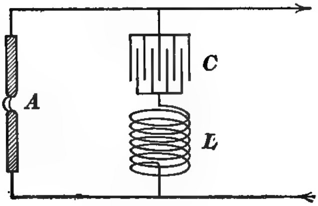


FIG. 227.

Just as irregular puffs of air blown over the mouth of an open cylinder set the inclosed air into regular vibrations, so the irregularities of the current flowing through the arc excite oscillations in the electric circuit. Currents flow in and out of the condenser, passing through the arc, alternately strengthening and weakening the previously constant current. This variation of the current, in its turn, affects the volume of the hot gases surrounding the arc, causing alternate expansions and contractions. Thus a compressional sound wave is sent out

from the arc, having the same vibration frequency as the electric oscillations.

404. Electrical Units. In this chapter we have used almost exclusively the *electrostatic system* of units, which was developed from the concept of unit charge, as given by Coulomb's law (Art. 362). From this fundamental unit all other units were derived, by using such relations as

$$V_1 - V_2 = \frac{W}{q}$$

$$I = \frac{q}{t}$$

$$C = \frac{q}{V_1 - V_2}$$

The *electromagnetic system* of units was developed from the fundamental concept of unit current (Art. 258), which was defined in terms of the magnetic effects of a current. From this all other units were derived by such relations as

$$Q = It$$

$$R = \frac{W}{I^2 t}$$

$$V_1 - V_2 = IR$$

While there is no difference in the nature of an electric charge, as defined in electrostatics, and of an electric quantity, as defined in current electricity, and no difference between the concepts of current, difference of potential, capacity, etc., as used in the two chapters, yet their units in the two systems are of very different magnitude. The quantity of electricity produced by electrostatic machines is always very much smaller than that produced by cells or other current generators, leading naturally to the selection of a smaller unit. In fact, one electromagnetic C. G. S. unit equals 3×10^{10} electrostatic units. On the other hand, the differences of potential produced by electrostatic machines are far larger than those produced by electric batteries or dynamos. One electrostatic unit of difference of potential equals 3×10^{10} electromagnetic units.

The factor of 3×10^{10} constantly recurs in the ratio between similar units in the two systems, and it may be shown to be equal to $\frac{1}{\sqrt{c\mu}}$, where c denotes *the dielectric constant* and μ *the permeability of the medium* in which the electrical phenomena occur. For a more detailed treatment of this subject, the student is referred to more advanced texts.

405. The Electromagnetic Theory of Light. By deriving the dimensional formulae of the various quantities, as used in the two systems of electrical units, it may be shown that the ratio $\frac{1}{\sqrt{c\mu}}$ has the dimensions of a velocity. Moreover, this is the velocity with which a periodic electrical disturbance is propagated through space. Very careful determinations of this ratio have fixed its value at $2.9971 \times 10^{10} \frac{\text{cm}}{\text{sec}}$.[1] But this is also the velocity of light. According to Maxwell's electromagnetic theory of light, which appeared in 1873, all ether radiations are considered as electromagnetic disturbances. He also predicted the existence of electrical waves, having all the properties of light waves. Fifteen years later Hertz (1857–1894) proved experimentally the truth of Maxwell's assumption by producing electrical waves and showing their identity with light waves (Art. 544).

Maxwell's theory has been further developed and modified by Lorentz and others, and has received very remarkable experimental verification in recent years (Art. 551).

The subject of electrical waves will be treated in connection with closely allied subjects in a separate chapter on Radiation.

Problems

[In this set of problems the answers involving electrical quantities are given in the electrostatic system of units, unless otherwise stated.]

1. Two small spheres, each weighing 0.1 g, and having equal charges, are suspended in air from the same point by silk fibers 80 cm long. If the spheres be kept 8 cm apart by repulsion, what is the charge on each?

Ans. 17.7 units.

[1] Rosa and Dorsey, *Bull. Bur. Standards*, vol. 3, p. 433, 1907.

2. Two small, equal balls having charges of + 10 and − 5 electrostatic units respectively are 5 cm apart in air. Find the force between them before and after contact with each other.

Ans. (*a*) Attraction, 2 dynes.
(*b*) Repulsion, 0.25 dyne.

3. A spherical conductor of 10 cm radius has a charge of 20 electrostatic units. Compute the surface density of the charge.

Ans. 0.016 unit per cm^2.

4. Compute the intensity of the electric field at a point 5 cm from a concentrated charge of 50 electrostatic units (*a*) in vacuo; (*b*) in a medium whose dielectric constant is 1.0005.

Ans. (*a*) 2 units.
(*b*) 1.9990 units.

5. Two small spheres 10 cm apart are charged with +5 and −5 electrostatic units respectively. Find the direction and magnitude of the field intensity at a point 10 cm from both charges.

Ans. 0.05 unit, parallel to direction from positive to negative charge.

6. Charges of 100, 200, 300 and 400 units are placed in this order at the corners of a square whose sides are 20 cm long. Find the direction and magnitude of the field intensity at the center.

Ans. 1.414 units, parallel to direction from 400 to 100.

7. Compute the field intensity at the center of the square in problem 6, when the charges are placed at the middle of the sides.

Ans. 2.828 units, parallel to diagonal.

8. Compute the potential at the center of the square in problem 6 (*a*) when the charges are placed as in problem 6; (*b*) when placed as in problem 7.

Ans. (*a*) 70.71 units.
(*b*) 100 units.

9. A conducting sphere of 5 cm radius is charged with +80 electrostatic units in air. Find the potential (*a*) of the sphere; (*b*) at a point 15 cm distant from its surface.

Ans. (*a*) 16 units.
(*b*) 4 units.

10. A spherical conductor of 10 cm radius is charged to a potential of 80 electrostatic units. What is the surface density of electricity upon the conductor?

Ans. 0.637 unit per $cm.^2$

11. An isolated conducting sphere of 10 cm radius having a charge of 40 units is connected by a long, thin wire to another isolated, uncharged conductor of 1 cm radius. Find the resulting potential of the two spheres.

Ans. 3.64 units.

12. Compute the capacity of a spherical condenser, the radii of the charged surfaces being 9.5 and 10 cm respectively, and the medium paraffine.

Ans. 380 units

13. Compute the capacities of two Leyden jars, whose tin-foil coverings have each an area of 200 cm^2, the thickness of the glass in one being 1 mm, in the other 2 mm ($c = 3$). *Ans.* (*a*) 477.5 units.
(*b*) 238.7 units.

14. Find the capacity of a plate condenser having on each side 10 plates, each 20×30 cm, separated by sheets of mica 0.1 mm thick.
Ans. 544,310 units.

15. What is the energy stored in the condenser of problem 14, when charged with 200 electrostatic units of electricity? *Ans.* 0.03675 erg.

16. Compute the amount of electrical energy disappearing when the two spheres of problem 11 are connected. *Ans.* 7.12 ergs.

17. What is the intensity of the electric field between two plane condenser plates which are 0.1 cm apart and differ in potential by 150 electrostatic units, the intervening medium being (*a*) air, (*b*) mica?
Ans. (*a*) 1500 units.
(*b*) 250 units.

18. Compute: (*a*) the capacity of the condenser of problem 12, p. 354, in electrostatic units; (*b*) the ratio between a microfarad and an electrostatic unit of capacity; (*c*) the ratio between an electromagnetic and an electrostatic unit of capacity.
Ans. (*a*) 1,910,000 units; (*b*) 1 microfarad $= 9 \times 10^5$ electrostatic units;
(*c*) 1 electromagnetic unit $= 9 \times 10^{20}$ electrostatic units.

THE ELECTRON THEORY

CHAPTER XLVI

ELECTROLYTIC CONDUCTION

*** 406. Early Theories.** The phenomena of electrolytic conduction (Chapter XXXV) are so different from those of metallic conduction that the theory proposed for their explanation grew up entirely independent of any other theory of electric conduction. In 1805, but a few years after the invention of the voltaic cell, von Grotthuss[1] laid the foundation of the theory which, in a modified form, is still held. According to his views, all molecules consist of positively and negatively charged atoms, held together by electrostatic attraction. In a solution the molecules are free to turn, and under the influence of a potential difference will place themselves in line with the electric field. If the difference of potential between the terminals of the electrolytic cell become sufficiently large, the molecules are torn apart into positive and negative particles, consisting either of atoms or groups of atoms. These charged constituents of the molecules are called *ions*. The positive ions, or *cations*, travel *with* the current, and the negative ions, or *anions*, travel *against* the current.

In 1857 Clausius[2] (1822–1888) modified this theory, in order to explain the fact that electrolytic decomposition may be obtained by very small differences of potential. According to Clausius, the collisions between the dissolved molecules and the water molecules are occasionally of sufficient violence to tear the molecules apart. Ions thus formed were assumed to be free for some time before recombining with ions having a charge of opposite sign. These free ions, which are always

[1] Von Grotthuss, *Mém. sur la décomposition de l'eau*, etc. Rome, 1805. *Also*, *Ann. Chim. Phys.* 58, p. 54, 1806.

[2] Clausius, *Pogg. Ann.* 101, p. 338, 1857.

present in an electrolyte, serve as carriers of electricity through the solution, even though the difference of potential applied to the terminals of the electrolytic cell be very small. When the ions reach the electrodes, they give up their electric charges, and the discharged atoms or groups of atoms either combine with each other, in obedience to some as yet unexplained chemical affinity, and form molecules, such as hydrogen or chlorine gas, *or* they act chemically upon the solution or upon the electrodes.

The charges of ions of the same kind are always the same, and hence equal masses of a given substance are decomposed by equal quantities of electricity (Art. 283). The charges of different ions are proportional to their valence (Art. 283). If we indicate the charge upon a univalent ion by a + or − sign, placed above the chemical symbol in each case, then the charges associated with ions of greater valence are represented by a number of + or − signs equal to the valence. Thus, common salt, NaCl, dissolved in water, is dissociated into $\overset{+}{\mathrm{Na}}$ and $\overset{-}{\mathrm{Cl}}$; silver nitrate, $AgNO_3$, into $\overset{+}{\mathrm{Ag}}$ and $\overset{-}{\mathrm{NO}}_3$; copper sulphate, $CuSO_4$, into $\overset{++}{\mathrm{Cu}}$ and $\overset{--}{\mathrm{SO}}_4$, and cuprous chloride, CuCl, into $\overset{+}{\mathrm{Cu}}$ and $\overset{-}{\mathrm{Cl}}$. The same quantity of electricity will thus liberate twice as much copper from a *cuprous* salt as from a *cupric* salt.

407. Electrolytic Dissociation Theory. After it had been found that all electrolytes, when dissolved in water, give abnormally large values for the osmotic pressure, for the lowering of the freezing point and for the raising of the boiling point, the theory of Clausius was further developed by Arrhenius[1] in 1887. It was shown that the number of molecules of the electrolyte dissociated on going into solution was, of necessity, much larger than Clausius had assumed. This theory, generally known as the *electrolytic dissociation theory*, is at present the leading theory of electrolytic conduction, and has not only led to a much better understanding of the phenomena concerned, but also to the discovery of important laws of electrochemistry.

[1] Arrhenius, *Ztschr. f. phys. Chem.* 1, p. 631, 1887.

According to this theory, the dissociation of the dissolved substances *increases with the dilution and becomes complete in very dilute solutions.* Only the ions are electrically and chemically active, while the undissociated molecules are inactive. Under the influence of a difference of potential, applied to the terminals of an electrolytic cell, the ions move through the solution with a definite velocity proportional to the potential gradient. The conductivity is directly proportional to the sum of the velocities of migration of the ions.

***408. Transfer of Electricity by Negative Charges.** From the point of view of the one-fluid theory (Art 377), we may assume only one kind of charge to exist. We shall assume, for reasons which will appear later, that it is the transfer of *negative* charges in a direction opposite to that of the current which gives rise to the phenomena of current electricity. Thus, the positive ions, upon reaching the cathode, do not give up a positive charge to the electrode, but take from it a negative charge, while the anions give up their charges at the other electrode.

409. Charge of an Ion. The electrolytic dissociation theory leads to the concept of very small but perfectly definite charges, which form the smallest quantities of electricity existing separately in electrolytic conduction. These are the charges carried by univalent ions. An atomic structure of electricity has been repeatedly advocated, and Weber, as early as 1871, called these charges "*atoms of electricity.*"

It becomes of interest to measure these charges. We have seen (Art. 284) that one chemical equivalent, for example, one gram of hydrogen, carries 96,530 coulombs. The number of ions in one gram of hydrogen may be found in the following manner. Various attempts have been made to determine from the theory of gases the number of molecules in one cubic centimeter of a gas. Loschmidt calculated this number for a gas at 0° C, and under a pressure of 760 mm of mercury, as 2.74×10^{19}, and Planck,[1] from thermodynamic reasoning, found 2.76×10^{19}. Now 2 g of hydrogen, that is, one grammolecule,

[1] Planck, *Ann. d. Phys.* 4, p. 564, 1901.

occupies under these conditions 22,390 cm^3. Hence, there are $11{,}195 \times 2.75 \times 10^{19}$ molecules of hydrogen in one gram. But each molecule of hydrogen consists of two atoms. Consequently, the number of atoms of hydrogen in one gram is

$$N = 11{,}195 \times 5.5 \times 10^{19} = 61{,}570 \times 10^{19} \tag{439}$$

These carry 96,530 coulombs; so each ion carries a charge,

$$\Delta q = \frac{96530}{61570 \times 10^{19}} = 1.57 \times 10^{-19} \text{ coulomb} \tag{440}$$

Since one coulomb equals 10^{-1} C. G. S. electromagnetic unit (Art. 263),

$$\Delta q = 1.57 \times 10^{-20} \text{ C. G. S. unit} \tag{441}$$

But one C. G. S. electromagnetic unit is 3×10^{10} as large as an electrostatic unit (Art. 404), and we obtain for the charge of a univalent ion

$$\Delta q = 4.7 \times 10^{-10} \text{ electrostatic unit} \tag{442}$$

Not a single experimental fact forces us to assume that these particles of electricity, when entering a metallic circuit, lose their individuality and combine to a continuous electrical fluid such as was assumed by the older theories. If we agree to take this more recent point of view, an electric current in a conductor may be nothing else than a transfer of such free separate charges through the spaces between the material atoms of the conductor. It would even be unnecessary to assume that these moving charges are connected with any ponderable matter, since the charges must have a separate existence, at least during the short time needed to pass from the ions to the electrodes of an electrolytic cell. In fact, recent discoveries in the field of electrical conduction through gases (Chapter XLVII) and radioactivity (Chapter XLVIII) strongly support such an interpretation of the phenomenon of an electric current.

CHAPTER XLVII

CONDUCTION THROUGH GASES

410. Influence of Pressure upon Discharge. If two metallic electrodes be sealed into the closed ends of a glass tube, about 50 cm long (Fig. 228), and connected to the terminals of a medium-sized induction coil or electrostatic machine, no discharge will occur through the tube so long as the air in the tube is under atmospheric pressure. If, however, the tube be exhausted, there soon appears, instead of the well-known spark, a discharge in form of a thin reddish line. Upon further exhaustion, the line begins to broaden, and at a pressure of about 1 cm of mercury

FIG. 228.

the luminous discharge nearly fills the entire tube. The cathode or negative electrode is covered with a layer of bluish light. Next to this is a darker space, called *the Faraday dark space*, and beyond this, extending to the anode, is a column of light of a reddish hue, called *the anode column* or *the positive column.*

Tubes of this kind present a splendid appearance, the color of the luminosity depending upon the nature of the inclosed gas. Fluorescent substances, such as uranium glass, kerosene or a solution of quinine, become beautifully luminous. Such tubes are frequently called *Geissler tubes.*

If the pressure be reduced still further, the tube changes in appearance. The positive column becomes less luminous, and breaks up into a series of light and dark layers, or striae. At a pressure of about 0.5 mm of mercury, the negative glow sepa-

rates from the cathode. At the same time a new, luminous layer develops at the cathode, separated from the first by a relatively dark space, *the Crookes dark space*, or *cathode dark space*. With still greater exhaustion the anode column practically disappears, and the cathode glow, while extending to a greater distance from the cathode, becomes weaker in luminosity, and at a pressure of about 0.01 mm disappears. The walls of the tube then begin to glow, usually with a bright greenish, fluorescent light.

411. Cathode Rays. The fluorescence of the walls of a highly evacuated tube is caused by a stream of very small particles, proceeding in straight lines from the cathode, and forming the so-called *cathode rays*. If the walls of the tube be protected from the impact of these rays, as, for example, by a thin sheet of metal placed inside the tube (Fig. 229), the shielded part of the glass will not become luminous. If after a short time the metal screen be removed, as, for example, by tipping the tube, so that the metal cross turns over to a horizontal position, on continuing the discharge, a bright cross upon a dimmer background will appear on the wall of the tube.

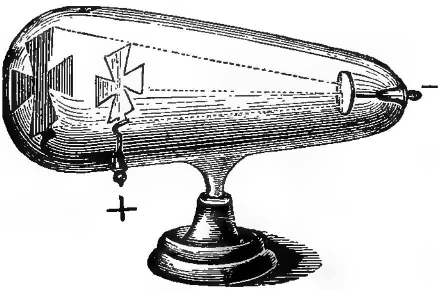


Fig. 229.

A delicately poised wheel with mica vanes will be set in rotation by the impact of these cathode rays, and will show by the direction of its rotation that the particles proceed from the cathode.

Cathode rays produce a marked heating effect when stopped. They excite many bodies to phosphorescence, and cause a change of color in some minerals. Their most important property is that they carry a negative charge. If a screen with a thin slit be placed in front of a cathode, a narrow beam of the rays passes through the slit. Its direction may be made visible by

placing behind the slit a phosphorescent screen. If now a magnet be brought near the tube (Fig. 230), the rays *are deflected in a direction* exactly *opposite to that in which a current would be deflected by the same magnetic field.* A deflection

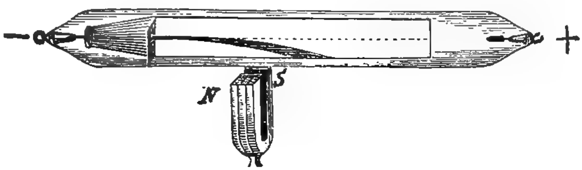


FIG. 230.

may also be obtained by placing such a tube in a strong electrostatic field.

***412. Lenard Rays.** After Hertz had shown that cathode rays are able to pass through very thin aluminium foil or gold leaf, Lenard investigated this phenomenon more thoroughly, and proved that the rays, after passing through the metal, retain all the characteristic properties of the cathode rays, although they can be detected but a very short distance beyond the thin metal window of the tube. These rays which have passed outside the cathode ray tube are often called Lenard rays, but are identical in their nature with the cathode rays.

***413. Velocity of Cathode Rays.** Suppose a charge e to travel with a velocity v in a direction at right angles to a magnetic field whose induction is B. The moving charge is equivalent to a current element of length l such that

$$Il = \frac{e}{t}l = ev \tag{443}$$

According to equation (404), a mechanical force F acts upon the moving charge, whose value is

$$F = BIl = Bev \tag{444}$$

This force, acting at right angles to v and B, produces a bending of the path of the particle in a plane perpendicular to v and B. As long as B remains constant, the deflecting force remains constant in magnitude, but is always directed at right

angles to the path. This is the condition of uniform circular motion, and the force may therefore also be expressed in terms of mechanical units, as

$$F = m\frac{v^2}{r} \tag{445}$$

where v is the speed of the particle, $\frac{1}{r}$ the resulting curvature of the path and m a measure of the kinetic reaction against the deflecting force, due to the inertia of the electromagnetic system. It does not necessarily follow from this that m must be ponderable mass. As already noted (Art. 329), an electromagnetic field shows effects similar to those due to the inertia of ponderable matter. We may call m the *electromagnetic mass of the charge*.

From equations (444) and (445) it follows directly, that

$$Bev = m\frac{v^2}{r} \tag{446}$$

or

$$Br = \frac{mv}{e} \tag{447}$$

Again, if the charged particle move at right angles to an electrostatic field of intensity E, it is deflected by a force F', whose value is given by equation (414), as

$$F' = Ee \tag{448}$$

If now a magnetic field and an electrostatic field be produced at the same time in the space through which the particle moves, and if the directions and intensities of these fields be adjusted in such a manner that the particle is not deflected under the influence of both forces, then evidently

$$Ee = Bev \tag{449}$$

and

$$v = \frac{E}{B} \tag{450}$$

If E and B be measured in the same system of units, their ratio gives directly the velocity of the charged particle. A number of experiments of this kind have been made on cathode rays, and it has been found that, while the velocity varies some-

what with the conditions of discharge, the velocity of the cathode rays is about 3×10^9 cm/sec, or one tenth of the velocity of light.

***414. The Ratio *e/m* in Cathode Rays.** It is also possible to calculate the ratio between the charge of the particles forming the cathode rays and their mass. Several methods have been employed for this purpose, all presenting great experimental difficulties. The following is theoretically very simple.

Let the particle receive its kinetic energy by passing through a difference of potential $V_1 - V_2$. The electrical energy expended is then $(V_1 - V_2)e$ and

$$(V_1 - V_2)e = \tfrac{1}{2} mv^2 \tag{451}$$

or

$$\frac{e}{m} = \frac{v^2}{2(V_1 - V_2)} \tag{452}$$

Combining the last equation with (450), we have

$$\frac{e}{m} = \frac{E^2}{2\,B^2}\frac{1}{V_1 - V_2} \tag{453}$$

All quantities on the right-hand side may be measured, and thus $\frac{e}{m}$ may be calculated. The best experimental results have given for the cathode rays

$$\frac{e}{m} = 5.1 \times 10^{17}\,\frac{\text{electrostatic units}}{\text{gram}}$$

$$= 1.7 \times 10^{7}\,\frac{\text{electromagnetic units}}{\text{gram}} \tag{454}$$

This value is independent of the manner in which the cathode rays are produced and of the nature of the metal forming the cathode.

415. The Electron. We have seen (Art. 284) that one gram of hydrogen ions carries 96,530 coulombs, or 9653 electromagnetic units. In this case the value for the ratio e/m is two

thousand times smaller than the corresponding value deduced from the cathode rays. Two explanations of this discrepancy may be considered. If the cathode particles are of a magnitude comparable with that of a hydrogen ion, their charges must be several thousand times larger than those of the ions. Or, we may assume that the charges are of the same magnitude as the ionic charges, but we are then forced to the conclusion that the mass of a cathode ray particle must be several thousand times smaller than that of a hydrogen atom. We shall see that, in the light of recent experimental results, this assumption appears to be the more reasonable.

These extremely small particles are called *electrons.* They are *negative charges.* It can be shown mathematically that the mass effect referred to (Art. 413) does not need to be due to ponderable matter connected with the charge. In fact, certain mathematical deductions require that the *mass of an electron shall increase with its velocity*, and this surprising conclusion has been verified experimentally. The fact that cathode rays are identical, *regardless of the source from which they may be derived, suggests that the electrons are common constituents of all atoms.*

Positive charges with masses comparable to that of an electron have not yet been found.

416. Canal Rays. In 1886 Goldstein,[1] while working with a discharge tube whose cathode was perforated by several holes, observed faintly luminous rays passing through the holes in a direction away from the anode. Where these rays met the wall of the tube, they excited a mauve-colored phosphorescence, totally different from that produced by cathode rays. These rays were called *canal rays.* Their direction indicated that they consisted of positively charged particles, but at first no experiments would give any indication of a charge. In 1898, however, Wien[2] showed that if sufficiently strong magnetic fields were employed, a deflection could be obtained in a direction

[1] Goldstein, *Berl. Ber.* 1886, p. 691.
[2] Wien, *Verh. d. Berl. phys. Ges.* 1897, p. 165.

opposite to that of the cathode rays. Subsequent measurements gave for these rays

$$v = 2 \times 10^8 \frac{\text{cm}}{\text{sec}}$$

$$\frac{e}{m} = 1 \times 10^4 \text{ electromagnetic units per gram}$$

The ratio of the charge to the mass is therefore the same as for hydrogen ions, and the velocity is independent of the potential difference between anode and cathode.

J. J. Thomson found, in 1907, that at very low pressures the particles in the canal rays were divided by the application of a strong magnetic field into two groups, and in helium gas a stage of exhaustion could be reached when a third well-defined group appeared. For the second and third group the values for $\frac{e}{m}$ were $\frac{1}{2}$ and $\frac{1}{4}$ of those found for the hydrogen ions. It is very probable that these canal rays consist of hydrogen and helium ions.

In 1910 J. J. Thomson[1] showed that the canal rays may be divided into three classes: —

(*a*) *Rays which are not affected by electric or magnetic fields.* Possibly these rays are formed by a recombination of negative and positive particles.

(*b*) *Secondary rays.* As the rays of the first type pass through the remaining gas and collide with the molecules, they produce these secondary rays. Whether they do this by splitting up themselves or by dissociating the molecules against which they strike, is uncertain, but the latter seems to be more probable. In ordinary discharge tubes these rays of the second class predominate and swamp the others. They are the rays described above, and are now assumed to originate in the space behind the cathode.

(*c*) *Rays characteristic of the gases in the tube.* These have been observed only at very low pressures and in large tubes. Their velocity depends upon the potential difference between the electrodes, and the value e/m for these rays is inversely

[1] J. J. Thomson, *Phil. Mag.* 20, p. 752, 1910.

proportional to the atomic mass of the gas from which they are derived. They originate between the anode and the cathode.

417. Roentgen Rays. In 1895 Roentgen[1] discovered that some sort of radiation, totally different from cathode rays, was produced outside of an ordinary cathode tube. These new rays, to which he gave the name X-*rays*, are now generally called *Roentgen rays*, after their discoverer. They are produced when cathode rays are suddenly stopped in their motion by striking a solid body. A very efficient form of Roentgen ray tube is shown in Fig. 231. The cathode concentrates the cathode rays upon a sheet of platinum, placed in the center of the tube. When the discharge of an induction coil or static machine is passed through the tube, the glass opposite the sheet of platinum shines with a bright green phosphorescence, and the presence of Roentgen rays outside the tube may easily be shown by their characteristic properties.

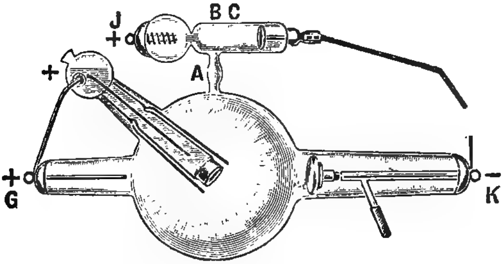


FIG. 231.

The exact nature of the Roentgen rays is not perfectly understood. Most physicists hold that they are pulses in the ether, propagated with enormous speed through space. They do not carry any electrical charges, and cannot be reflected or refracted as light waves are.

418. Properties of Roentgen Rays. (*a*) *Roentgen rays excite phosphorescence* in a large variety of substances, such as the double sulphate of potassium and uranium, crystals of willemite or of platinocyanide of barium. A screen of cardboard, covered with a thin coating of any one of these substances,

[1] Roentgen, *Würzb. Ber.* 1895, p. 137.

FIG. 232.

shines with characteristic phosphorescence when placed in the path of Roentgen rays.

(*b*) *These rays have great penetrating power, being able to*

pass through bodies of considerable thickness. Different substances absorb Roentgen rays in different degree, as is well illustrated in the case of the parts of the human body. If the hand be placed on the back of the phosphorescent screen between the Roentgen ray tube and the screen, a distinct shadow picture or silhouette will be seen upon the screen (Fig. 232). The bones absorb the rays more strongly than the fleshy parts, and the shadow cast by the bones appears dark upon a lighter background. Metals absorb these rays quite strongly, though some rays are able to penetrate a lead sheet a few millimeters thick.

(*c*) *Roentgen rays produce photographic action similar to that due to light.* Since the rays pass easily through wood or hard rubber, Roentgen ray photographs may be taken without removing the cover of the plate holder. It should be kept in mind that these so-called photographs are not obtained by reflection from the bodies, but are merely silhouettes or shadow pictures of the bodies through which the rays pass.

(*d*) *Gases through which Roentgen rays pass become conductors of electricity.* Thus, if a charged electroscope be placed in the neighborhood of an active Roentgen ray tube, it will be found that the gold leaves collapse, since the charge of the instrument is rapidly carried away by the conducting air.

419. Ionization of Gases. The electrical conductivity of gases in their normal state and under atmospheric pressure is extremely small. But when Roentgen rays are passed through a gas, its conductivity increases enormously. When the rays cease to act, the conductivity disappears in a short time. The theory offering the best explanation of these phenomena is fashioned after the electrolytic dissociation theory. According to this theory, positive and negative ions are produced in a gas which is exposed to the action of Roentgen rays. It has, however, not been proven that these ions consist of particles smaller than molecules. In fact, it has been found that frequently a number of molecules are clustered about a charge, forming ions of relatively large mass. The mass of these ions is variable.

As soon as the Roentgen rays cease, the ions recombine, and neutral molecules are formed. This recombination does not take place instantly, but the ionization persists for some seconds. Thus, ionized gas may be drawn through a tube and still retain its power of discharging an electroscope, though it has been removed from the influence of the Roentgen rays. If, however, a cotton plug be placed in the tube, the conductivity of the gas is entirely destroyed.

It is to be noted that recombination of the gaseous ions occurs not only upon the cessation of the Roentgen rays, but takes place during their action as well. This is shown by the fact that, in any mass of gas subjected to the action of Roentgen rays, a definite state of equilibrium between ionization and recombination always occurs.

If a gas placed between two metal plates be ionized and a difference of potential be established between the plates, the positively charged ions travel toward the lower potential and the negatively charged ions toward the higher potential. This is equivalent to an electric current passing between the plates, and the current, though in general very small, may be measured by a sensitive galvanometer.

*420. **Other Sources of Ionization.** Ionization of gases may be produced by other means than Roentgen rays. For example, gases in the neighborhood of incandescent bodies conduct fairly well. The gases of a flame always exhibit high conductivity, and will rapidly discharge electrically charged conductors. Ultra-violet light is an efficient ionizer, and the discharge of a condenser will take place at a much lower potential difference when the spark gap is illuminated by ultra-violet light than in diffused light. The effect of these ionizing influences is much weaker than that of Roentgen rays. The radiations from radioactive substances, which will be treated in the next chapter, are the most powerful ionizers known, and are at present used almost exclusively in the study of ionization of gases.

*421. **Ions as Nuclei.** We have seen (Art. 215) that dust particles form the nuclei around which condensation of water

vapor begins. Dust-free air, however, must be cooled considerably below the dew point before the water vapor contained in it will condense. But Wilson found that it requires much less supercooling to produce condensation of water vapor in dust-free air if the air be ionized, and that the formation of drops begins at an earlier stage around the negative ions than around the positive ions.

Supercooling of a volume of gas containing moisture may be produced by sudden expansion. Thus, let the ionized gas be inclosed in a vessel. By a proper adjustment of the amount of expansion, the droplets may be made to form only around the negative ions or around both kinds, as may be desired. If a large number of ions be present, a fine mist will be formed, which slowly sinks to the bottom of the vessel under the action of gravity.

*** 422. Charge of an Ion.** If the expansion be so regulated that the condensation takes place only around the negative ions, the total charge carried by these ions will be transferred by the drops to the bottom of the vessel, and may be measured by a sensitive instrument. It is also possible to calculate the diameter and consequently the mass of the individual droplets from the rate at which they sink through the air. If then the whole mass of the condensed water be measured and be divided by the mass of a single drop, the total number of drops, that is, the total number of ions present, is found at once. Dividing the total electric charge by this number, the charge upon each individual ion is obtained. Experiments of this kind gave about 4×10^{-10} electrostatic unit as the charge upon each ion. On account of the evaporation of the water, this method presents great experimental difficulties.

Recently Millikan[1] has modified this method by blowing a cloud of very fine droplets of oil by means of an atomizer over a horizontal air condenser and allowing a few droplets to enter the space between the horizontal plates of the condenser. These droplets sink slowly through the air under the action of

[1] Millikan, *Science*, 32, p. 436, 1910.

gravity, and their rate of fall may be measured by means of a telescope focused upon an individual droplet. If now the plates of the condenser be charged to a certain difference of potential, the rate of descent of the droplet will not be affected unless it possess a charge. In fact, the droplets were always found to be charged on entering the observation chamber. This charge was probably due to friction in the nozzle of the atomizer.

Now the difference of potential between the plates may be so adjusted that the force on the charged droplet due to the action of the electrostatic field nearly neutralizes the effect of gravity, and the droplet may be kept under observation for a long time. During his experiments Millikan found that a droplet frequently caught or lost one or more ions, which resulted in an immediate change in its motion.

With the electrical field cut off, the droplet was observed while falling under the action of gravity through a definite distance, and the time required was noted. Then the field was thrown on, and under its influence the droplet moved upward. Again the time was noted during which the drop passed over the same distance as before.

Now it may be shown that under the conditions of the experiment the speed of the droplet is proportional to the forces acting upon it. If v_1 be the speed under the action of gravity and v_2 the speed resulting from the combined action of gravity and of the electrical field of intensity E, the following relation holds:

$$\frac{v_1}{v_2} = \frac{mg}{Ee - mg} \tag{455}$$

or

$$e = \frac{mg}{Ev_1}(v_1 + v_2) \tag{456}$$

By this ingenious method Millikan was able to calculate the charge of an ion with great precision.

Observations have also been made by a number of other methods. The best value of the charge of an electron is now believed to be

$$e = 4.65 \times 10^{-10} \text{ electrostatic unit.}$$

***423. Charge of an Electron.** We found (Art. 409) that the smallest electric charge, taking part in electrolytic conduction, is the charge of a univalent ion. This was calculated to be 4.7×10^{-10} electrostatic unit. The study of ionization of gases also leads to a definite elementary electric charge of the same magnitude, which, therefore, may be justly called *an atom of electricity*. Since these charges are always observed either singly or in very small multiples, we are justified in the assumption that the *charge of an electron is this elementary charge of negative sign*, and that therefore the mass of an electron is very small, or only a minute fraction of the mass of a hydrogen atom.

***424. Applications of the Electron Theory.** From the point of view of the electron theory, electricity is of one kind only, namely, negative electricity. A negatively charged conductor should no longer be thought of as being covered uniformly over its whole surface with electricity, but as having attached to it a large number of separate electrons. The properties of a positively charged body are to be considered as mainly due to a loss of electrons.

The electrons are of much greater mobility under the influence of an electric field than the heavy, positively charged particles, and a current must therefore be considered as being due mainly to the transference of electrons, though their direction is of course in the opposite sense to that of the current, as defined in previous articles. The electron theory thus shows a marked similarity to Franklin's one-fluid theory.

It should, however, be kept in mind that the electron theory does not necessarily mean a return to the action-at-a-distance theory. The transfer of an electron is always accompanied by a disturbance in the medium about the conductor, as shown by the phenomena of electromagnetic induction. A very close connection must therefore exist between the electron and the medium, but the nature of such connection is at present unknown.

CHAPTER XLVIII

RADIOACTIVITY

425. Discovery of Radioactivity. In 1896, just after the discovery of the Roentgen rays, Becquerel[1] (1852–1908) investigated the action of various phosphorescent substances upon a photographic plate, believing that the emission of Roentgen rays was connected with the green phosphorescence of the glass wall of the tube. None of the substances investigated had any effect, except uranium salts, but he also found that their action was entirely independent of any phosphorescence, for the effect persisted long after all phosphorescence had disappeared. Becquerel established the fact that uranium salts emit rays which in many respects are similar to Roentgen rays, and which were at first called Becquerel rays, after their discoverer. But they were soon found to be a mixture of three different kinds of rays, which are now called the α, β and γ rays.

About a year after Becquerel's discovery it was found that thorium salts possessed the same property as the salts of uranium. Substances which emit Becquerel rays are said to be *radioactive*. Great progress in this field was made when M. and Mme. Curie[2] succeeded in separating from pitchblende certain bismuth salts whose radioactive power was about 400 times that of uranium. The active substance in these bismuth salts was called *polonium*. Soon after, they, in conjunction with Bémont, succeeded in separating from pitchblende the chloride of a new element, *radium*, which shows very powerful radioactive properties. Another radioactive substance, which is found in thorium minerals, was discovered by Debierne in 1899, and was called *actinium*. The chemistry of the radioac-

[1] Becquerel, *C. R.* 122, p. 301, 1896.
[2] Curie, *C. R.* 127, pp. 175, 1215, 1898.

tive substances is still unsolved, but radium has been proved beyond doubt to be an element, with a characteristic spectrum. The amount of this element which can be obtained is exceedingly small, since from a ton of pitchblende only a few milligrams of radium chloride can be separated. In 1910 Madame Curie and Debierne succeeded in obtaining radium in the metallic state.

426. Properties of the Radiations. All radioactive substances send out radiations with the following properties. A photographic plate is affected, even if it be protected from light. The radiation produces phosphorescence, ionizes gases and discharges charged conductors. No reflection, refraction or polarization of these rays has ever been observed. The most important property, from a theoretical point of view, is that at least a portion of the rays are deflected by a magnetic field, and this led to the discovery that the rays are not homogeneous, but consist of three kinds of rays, which are called the α, β and γ rays respectively.

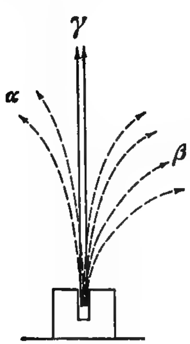


FIG. 233.

427. The α Rays. When the rays from a radioactive substance are made to pass normally through a magnetic field, a portion of the rays is deflected towards one side, another group in the opposite direction, while a third group is not deflected (Fig. 233). The first group consists of positively charged rays which are called α rays. These are the least penetrating of Becquerel rays, since no α rays are known to penetrate 10 cm of air under atmospheric pressure or through a couple of sheets of note paper, without losing their ionizing property. On the other hand, they produce nearly all the ionization of a gas, exposed directly to Becquerel rays.

Rutherford found that $\frac{e}{m}$ for these rays is the same, from whatever element the rays are emitted, and that its value is 0.5×10^4 electromagnetic units per gram, or the same value which was found for one of the three groups of the canal rays

(Art. 416). The velocity of these rays is never very different from 2×10^9 cm per second. In fact, they differ from the canal rays only in their greater velocity. The α rays produce the spectrum of helium, and have been proven by Rutherford to be identical with *positively charged helium atoms.*

428. The β Rays. The β rays have a negative charge, and in general are much more penetrating than the α rays, although their penetrating power varies within wide limits. Some appear no more penetrating than α rays, while others are able to produce ionization after passing through half a centimeter of lead. Most photographic action is due to the β rays. They are easily deflected by a magnetic field, and the values for $\frac{e}{m}$ and v are almost identical with those found for cathode rays. They are therefore *electrons.* Some of these rays have greater velocities than cathode rays. Thus, Kaufmann observed velocities as high as $2.85 \times 10^{10} \frac{\text{cm}}{\text{sec}}$, which is nearly the velocity of light. When the velocity falls below $3.6 \times 10^8 \frac{\text{cm}}{\text{sec}}$, they are unable to ionize a gas. Rays with a smaller velocity than this have been observed by Thomson by means of the charge which they carry; it has been proposed to call these slowly moving electrons δ rays.

429. The γ Rays. The third group of Becquerel rays are called γ rays. They are more penetrating than β rays, and produce ionization even through several centimeters of lead. They are not deflected by the strongest magnetic fields which may be produced experimentally. It is now generally held that they are *Roentgen rays*, and consist of electromagnetic pulses propagated with great velocity through space.

*** 430. Radioactive Energy.** It was first shown by the Curies that the temperature of radium salts is always several degrees higher than that of the surrounding bodies. Since heat is continually conducted away and radiated from the vessel in which the radioactive substances are kept, the maintenance of a

higher temperature indicates that energy (in the form of heat) is constantly given out by radioactive substances.

The experiments of St. Meyer and Hess in 1912 and others show that one gram of radium emits heat at the rate of about 132 calories per hour, or

$$\frac{H}{t}=\frac{135\times 4.2\times 10^{7}}{3600}=1.5\times 10^{6}\,\frac{\text{erg}}{\text{sec}}$$

***431. Theory of Radioactivity.** The theory of radioactivity accepted by most physicists is that proposed by Rutherford and Soddy.[1] In accordance with this theory, radioactive phenomena are due to a continuous disintegration of the radioactive substance. In 1900 Crookes, by chemical means, separated from uranium a substance which seemed to contain all the radioactivity of the uranium, while the remaining uranium showed no activity whatever. But further experiments have shown that the apparently inactive uranium still retained the power of sending out α rays, but no β rays. However, when Crookes examined this uranium after the lapse of a year, it had completely regained its power to emit β rays, and again a substance could be separated from it which produced β rays, while the remaining uranium did not do so.

It is therefore clear that uranium, when left to itself, undergoes a change which consists in the formation of *another substance*, which has the power of producing β rays, and which is called uranium X. The conclusion seems justified that we have here a change in the *atom of uranium itself, or a transformation of one element into another element.*

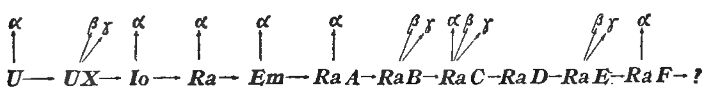
***432. Decay of Radioactive Substances.** The disintegration theory explains also why radioactive substances produce large amounts of heat (Art. 430). This heat is simply the equivalent of the difference of the internal energies of the atoms before and after transformation. We must further expect that the original radioactive substance will disappear in course of time.

An immense amount of work has been done to measure the

[1] Rutherford and Soddy, *Phil. Mag.* 4, p. 370, 1902.

average life of a radioactive atom, and the results show that this time varies for the different substances from six hundred million years for uranium to three seconds for actinium emanation, which is a radioactive substance obtained from actinium.

The study of the products of disintegration of the radioactive substances has led to the discovery of many consecutive products, differing from each other in their chemical nature, disintegration period and the kind of rays which they emit. Rutherford has worked out a complete series of the products of radium, which itself is probably a disintegration product of uranium, though not the first. The following sketch represents this series, starting with uranium, *U*. *UX* denotes uranium X, *Io* ionium, *Ra* radium, *Em* the first product of radium, called the *radium emanation*. The arrows indicate the kind of rays produced by each of the substances. Similar series have been worked out for other radioactive substances, such as thorium and actinium.



Almost all the products give off α rays, except radium D, which is not radioactive, and uranium X, radium B and radium E, which produce only β and γ rays. Radium C sends out all three kinds of rays. The radium emanation is a gas, and has been liquefied at $-150°$ C. It belongs to the argon family. Radium C is a solid at ordinary temperatures, and radium F is polonium, the first radioactive substance separated from pitchblende by the Curies. Radium F is transformed into a substance which has no radioactive properties, and is at present unknown, but there is a strong belief among physicists that this last product is lead, which is always found together with radium and helium in uranium minerals.

More recent investigations have shown that uranium is probably a mixture of two radioactive substances, namely uranium 1 and uranium 2; and that uranium 2 may give rise to a branch product uranium Y which decays to half its amount in

1.5 days. There is at present no definite information whether or not uranium Y gives rise to other successive products and thus forms a series parallel with the products from uranium X. It has been shown further that radium C is of a complex nature and gives rise, not only to radium D, but also to a short-lived product, radium C_2 which emits β rays. Only about the $\frac{1}{6000}$th part of the disintegration product of radium C appears as this branch product C_2. Though very little is known about uranium Y and radium C_2, their discovery is of great importance since it tends to show that radioactive atoms may break up in at least two different ways.

The atomic mass of uranium is 238.5, that of ionium 230.7, that of radium 226.6 and that of lead 206.5. The differences between these atomic masses, viz. 7.8, 4.1 and 20.1, are very nearly simple multiples of 4, the atomic mass of helium, which in the form of α rays is a by-product of this process of disintegration. This, in connection with the fact that the canal rays consist, at least in part, of charged helium atoms, points to these atoms as elementary, stable complexes which serve as building stones for the more complicated atoms. The electron theory may be able to furnish an explanation of the nature of an atom, but a discussion of this problem does not come within the scope of an elementary textbook.

LIGHT

INTRODUCTION

CHAPTER XLIX

FUNDAMENTAL PHENOMENA

433. Definitions. Optics is that branch of physics which has for its object the study of the nature of light and the circumstances of its propagation. In *Geometrical Optics*, the circumstances of the transmission of light are deduced from certain laws established by experiment. These laws are: (*a*) the law of the rectilinear propagation of light; (*b*) the law of the independence of the different portions of a beam of light; (*c*) the law of reflection; (*d*) the law of refraction. In *Physical Optics* these laws are explained in accordance with certain assumptions regarding the constitution of matter and the nature of the space which it occupies. Geometrical optics makes no assumptions regarding the *nature* of light, but through the application of the above geometrical laws which are true, whatever the nature of light may be, it deduces important formulae and explains many optical phenomena. Physical optics, on the other hand, explains many phenomena that cannot be accounted for upon geometrical principles, and anticipates in many cases the results of experiment.

A body which emits light of itself, as the sun, the fixed stars, lamps, etc., is said to be *self-luminous*. Bodies which emit no light of themselves, but which shine by reflected light, are called *non-luminous* bodies.

If a body transmit light freely, it is said to be *transparent*. If it transmit light but poorly, so that the outlines of a body cannot be seen through it, it is called *translucent*. An *opaque* body

is one that transmits no light. As usual in physical properties these distinctions are not absolute. No substance is perfectly transparent, neither is any entirely opaque.

A *ray* of light denotes nothing more than the direction along which the light is propagated. A symmetrical collection of rays about some axis is called a *pencil.* If the rays meet at a point, such a point is called the *focus* of the pencil.

Any space or substance in which light can be propagated is called a *medium.* Only isotropic media are considered in geometrical optics. In the study of the transmission of light in different media, the plane in which the ray is assumed to lie is the plane containing the normal to the surface of the medium. The angle of *incidence*, of *reflection* or of *refraction* is the angle included between the *normal* and the *incident*, *reflected* or *refracted* ray. In isotropic media these angles all lie in the same plane, that is, *in the plane of incidence, which is defined by the incident ray and the normal.*

434. Nature of Light. The most casual observation shows that when light is transmitted from one body to another, energy is transmitted at the same time. Living plants and animals are stimulated and influenced by the action of light. Certain chemical reactions are induced by light; a mixture of hydrogen and chlorine in equal volumes explodes on exposure to the sunlight.

Again, light is developed only at the expense of energy in some other system, as in the falling stars and meteors, in the filament of an incandescent lamp, in the flame of a candle or in the faint glow of the firefly. The chief source of all terrestrial energy is the sun. The energy of the sun requires a little over eight minutes to traverse the gulf of intervening space between the sun and the earth. The question arises, where is the energy during that interval of eight minutes? Now since the only way in which energy can be transferred from one body to another remote from it, is either by the bodily transference of matter, as in the case of a projectile, or by disturbing the medium surrounding the second body, the answer to the foregoing question necessitates a choice between two assumptions. Either we are to assume that light is transmitted across space in the form of

small particles which carry the energy with them, or we must assume the existence of a continuous medium in space through which light is transmitted in some form of wave motion. No third mode of transferring energy is mechanically conceivable.

The first of these assumptions lay at the bottom of the corpuscular or emission theory as elaborated by Newton. The assumption of an hypothetical medium underlies the undulatory theory as founded by Huygens and developed by Young and Fresnel. This hypothetical medium is called the ether. It is assumed to fill all space and to permeate all bodies, to be of enormous elasticity, to offer no resistance to the passage of bodies through it, yet to be capable of resisting a shearing stress. According to the more modern view (Art. 405) the ether is subject only to electromagnetic stresses.

The undulatory theory of the propagation of light in some form or other is generally accepted to-day, and although in the articles under geometrical optics no theory of propagation is mentioned, since none is needed, it is to be understood as being tacitly assumed. As a working hypothesis we may assume that light consists in a periodic change of condition which is propagated with finite velocity in the form of transverse waves. While there is nothing in the geometrical treatment inconsistent with the undulatory theory, yet regard for simplicity has suggested that its formal statement be deferred until a later chapter.

435. Rectilinear Propagation. One of the earliest facts of observation upon optics is that light in a homogeneous medium appears to travel in straight lines. Opaque bodies are illuminated only on the side turned toward the source of light. It is impossible to see through a tube bent so that no straight line can be passed through it. The mechanic, the surveyor and the astronomer alike recognize and use this fundamental fact.

A closer examination of the phenomena will show, however, that rectilinear propagation of light is not only confined to homogeneous media, but is only approximately true even there. It will be shown later that owing to its fundamental nature, light does bend round corners, and the edges of shadows of

small objects are not of the size and sharpness demanded by rectilinear propagation.

In accordance with the principle of Huygens (Art. 116), the transmission of light in straight lines is a direct consequence of the interference of the subsidiary waves, which start out from every point upon the wave front. So long as the physical properties of the medium are the same in all directions, the wave fronts are spherical in form, and transmission of the luminous disturbance takes place in right lines normal to the wave front. In media of varying density, or of unequal elasticity, the wave form is no longer spherical, and the velocity of the disturbance is different in different directions. In media in which the density varies gradually from point to point, the transmission of light may even take place along curved lines. This is exemplified in the distortion of images of objects seen through air rising from heated surfaces, in the effects of mirage, and in the twinkling of stars.

436. Shadows. A result of the rectilinear propagation of light is the formation of dark spaces in the rear of opaque objects of any size when exposed to a source of light. The dark outline of an opaque body projected upon a screen by a luminous point is called its shadow. Such an outline may be found by drawing straight lines from the luminous point to the screen past every point of the exposed surface of the body. The figure obtained is termed the *geometrical shadow*. If the source of light be very nearly a point, then all the light is cut off from the screen, over a certain space, which is called the *umbra*. If the luminous body have any dimensions (Fig. 234), there will always be a ring of *partial shadow* surrounding the umbra. This is called the *penumbra*.

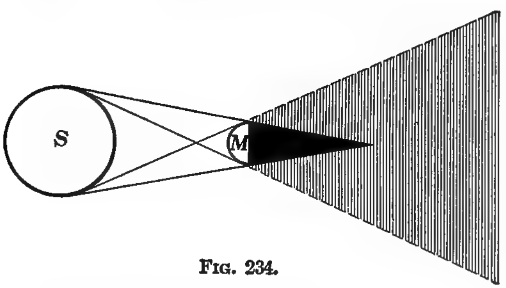


Fig. 234.

If the luminous body be larger than the opaque body, the umbra has a definite length. In the case of spheres the umbra becomes a cone of shade extending out into space, surrounded by a region of penumbra which increases with the distance from the opaque body. In eclipses of the moon the moon passes through the cone of earth shadow, entering the penumbra first and leaving it last. In solar eclipses the tip of the cone of shadow from the moon may or may not extend to the earth, according to the position of the moon in its orbit, since the mean length of the lunar shadow is less than the mean distance from the moon to the earth. If the shadow cone reach the earth, the solar eclipse is *total* for all points touched by the umbra; for points touched by the penumbra only, the elipse is *partial*. If the tip of the cone fall short of the earth, the eclipse is *annular* for the successive points touched by the prolongation of the axis of the cone of shadow.

437. Images through Small Apertures. An *image* may be defined as the picture of an object, formed by the crossing, either real or apparent, of a series of rays from the various points of an object. If the rays actually pass through the image, it is a *real image*, and may be caught upon the hand or projected upon a screen. If the rays *seem* to come through the image but do not, the image is *virtual*. If a luminous body *CD* (Fig. 235) be placed in front of a small aperture *H*, in an opaque screen *A*, there will be depicted upon a second screen *B* an inverted image of the body. This arises from the rectilinear propagation of light. Each point on the luminous body acts as an independent source of light, from which a cone of rays pass out toward the screen.

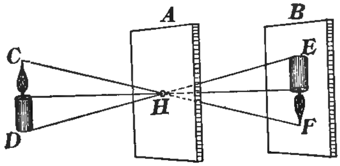


FIG. 235.

Since the only rays from the extremities of the object that can reach the screen *B* are *CH* and *DH*, it follows that all the light reaching the screen from the point *C* will be collected at *F* on the image, and similarly for the point *D* and its image at *E*. The position of the image is consequently *inverted*, since

all the rays forming the image cross at the aperture. The image, if viewed from the side toward the aperture, is also ***perverted***, *i.e.* the parts of the image relative to the object are symmetrically displaced with regard to a vertical plane through its center and normal to *B*. A glove for the right hand is the perverted image of that for the left hand. No amount of rotation can make them agree. An image seen in a plane mirror is perverted.

The *distinctness* of the image has reference to the sharpness of its outline. If the aperture *H* be made smaller, fewer rays from the point *C* would reach *F*, but these would all be from *C*, hence there would be no overlapping of images. The *distinctness*, therefore, *decreases as the size of the aperture increases.*

The *brightness* of the image refers to the amount of light per unit area that reaches the screen *B*. Manifestly the brightness will be greater the larger the aperture, or the brightness *varies directly as the size of the aperture.*

The *size* of the image may be determined from geometrical considerations. The two triangles *CHD* and *EHF* are similar, and we have at once the *law of linear dimensions;* that is, *the size of the image is to the size of the object directly as their distances from the aperture.*

It remains to be noted that the conditions for *brightness* and *distinctness* are mutually opposed, and that with a simple aperture one must be sacrificed to obtain the other. A converging lens placed in the aperture will secure both *brightness* and *distinctness* at the same time.

GEOMETRICAL OPTICS

CHAPTER L

PHENOMENA OF REFLECTION

438. Reflection of Light. When light, traveling in one homogeneous medium, meets the boundary of another homogeneous medium of different optical property, it is in general divided into several parts which follow different paths. A part is *reflected*, or turned back into the first medium along definite lines. A part is *scattered*, or reflected in all directions, a part is *transmitted* along one or two new paths, and a part is *absorbed*.

Scattering of light is produced by reflection at irregular or rough surfaces. All non-luminous bodies are rendered visible by scattered or *diffused* light. Even the most highly polished mirrors scatter some light. In general the quantity of light reflected from a surface increases as the angle of incidence increases. In one case, to be discussed later, the reflection is *total*.

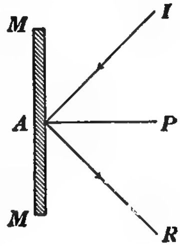


FIG. 236.

When a beam of light is allowed to strike the surface of a piece of highly polished glass or metal, the light leaves the surface along definite lines and is said to be reflected, as in Fig. 236, which shows the reflection of light by a plane mirror. The angle IAP is the *angle of incidence*, and the angle PAR is the *angle of reflection*. The law of reflection of light states that *the angle of incidence is equal to the angle of reflection, and the two angles lie in the same plane.*

The experimental proof of this law lies in the fact that with

the best mirrors and circles which men have been able to make, no one has ever found the slightest deviation from its truth.

439. Images in a Plane Mirror. When an illuminated object is placed before a plane mirror, rays of light pass off from it to the mirror and are reflected to the eye. An image of the object is seen reflected in the mirror, in the direction *from which the light enters the eye.*

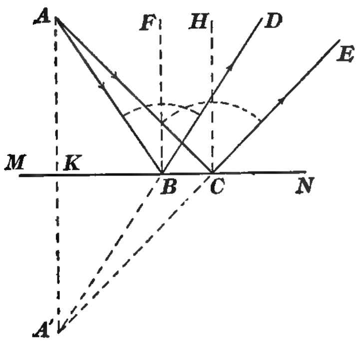


FIG. 237.

In accordance with the law of reflection, any two rays AB and AC (Fig. 237) will, if produced backward after reflection, seem to meet along the lines DB and EC, as if they originally came from some point as A_2 (not shown in figure), and an eye so placed as to receive these two rays would seem to see the point A at A_2.

The position of A_2 is readily found. From A drop a normal to the mirror cutting it in K, and extend it to meet BD produced backward in A'. Then the angles

$$ABM = DBN = KBA', \tag{457}$$

and the angles at K are right angles, hence the triangles AKB and $A'KB$ having the side KB common and the three angles equal each to each, are similar and equal. Therefore

$$AK = KA'.$$

Next draw the line $A'C$. Then the two triangles AKC and $A'KC$, since they have the two sides AK and $A'K$ equal, the side KC common, and the angles at K right angles, are equal in all their parts.

From this it follows that

$$ACK = A'CK = ECN \tag{458}$$

or the line $A'CE$ is a straight line.

Hence we see that the points A' and A_2 coincide, and that the

image seen in a plane mirror lies as far behind the mirror as the object lies in front of it.

Again, since the angles BAC and $BA'C$ are equal, the divergence of the rays is not changed by reflection in a plane mirror, and consequently the image is not distorted.

440. Path of Rays. The image of an object seen in a plane mirror may be constructed, and the paths of the individual rays determined graphically, in accordance with the foregoing principles. Let AB (Fig. 238) represent an object in front of a plane mirror MN. If the positions of the images of A and B can be determined, the complete image may be obtained by joining these two points. From A and B drop perpendiculars upon the mirror and produce them as far behind the mirror as the points A and B lie in front of it. The positions of the points A' and B' are thus determined. To an eye situated at E or E' rays seeming to come from A' really touch the mirror at the points where a straight line from E or E' to A' intersects the mirror, and since the rays originally came from A, these points are to be joined to A. The paths of the rays from B are found in the same way.

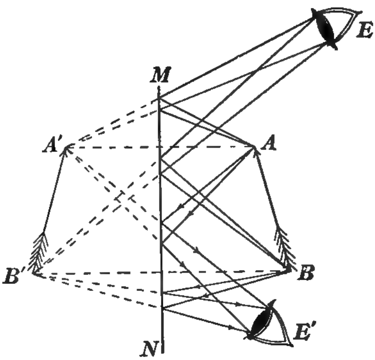


Fig. 238.

441. Deviation produced by Rotation of Plane Mirror. If a ray of light fall upon a plane mirror, and the mirror be turned through any angle, the reflected ray will be turned through twice that angle. For let AM (Fig. 239) be a ray of light incident normally upon the mirror M. The ray after reflection retraces its path, since the angles of incidence and reflection are both zero. When the mirror is turned through an angle θ, the normal to the mirror is rotated through the same angle.

The angles of incidence and reflection are now each equal to θ while the total deviation of the ray is $2\,\theta$.

This principle finds wide application in physical apparatus where it is desired to detect small angular movements or measure the same with great accuracy. Examples of such applications are seen in the rotating mirror as applied to the examination of manometric flames; to the measurement of the velocity of light; in the reflecting galvanometer, where a small mirror attached to the moving system indicates very slight angular displacements; and in the optical lever, where it is applied to the measurement of small lengths.

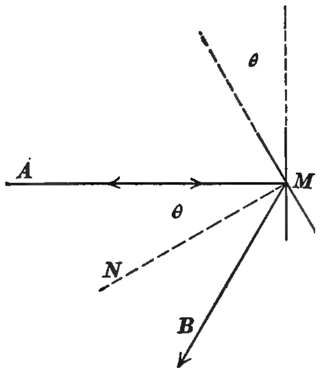


FIG. 239.

442. Successive Reflection from Two Mirrors. Let AO and BO (Fig. 240) represent two plane mirrors inclined to each other at an angle ϕ, and let $PQRST\ldots$ be a ray of light

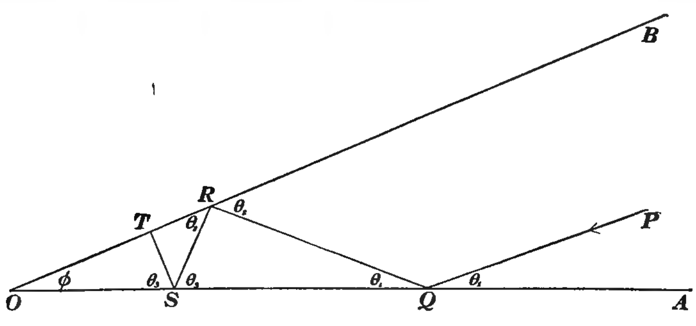


FIG. 240.

which is reflected in turn at Q, R, S, $T\ldots$. Let θ_1, θ_2, $\theta_3\ldots$ be the acute angles included between the ray and the mirror surfaces at the successive reflections. Then from the triangle QOR, we have

$$\theta_2 = \theta_1 + \phi$$

and from RSO we get
$$\theta_3 = \theta_2 + \phi \tag{459}$$
and so on.

These equations may be written
$$\begin{aligned} \theta_2 - \theta_1 &= \phi \\ \theta_3 - \theta_2 &= \phi \\ \cdot \quad \cdot \quad \cdot & \quad \cdot \quad \cdot \quad \cdot \\ \theta_{n+1} - \theta_n &= \phi \end{aligned} \tag{460}$$
whence by addition
$$\theta_{n+1} - \theta_1 = n\phi \tag{461}$$

When n is even, the angles θ_{n+1} and θ_1 are measured from the same mirror and the difference is the angle between the initial and final directions of the ray. Hence in this case *the total deviation is equal to n times the angle included between the two mirrors.*

Again, the deviation is the same whatever be the angle of incidence, so that the divergence of any two rays is not changed by successive reflections.

When the ray is twice reflected, once at each mirror, the total deviation is twice the angle between the two mirrors. This principle is applied in the sextant, an instrument invented by Newton and constructed by Hadley for measuring the angular distance between two remote objects. By means of the sextant the mariner is able to determine his latitude and longitude at any time when the sun is visible.[1]

443. Concave Spherical Mirrors. A concave spherical mirror (Fig. 241) is a part of a spherical shell with its inner surface

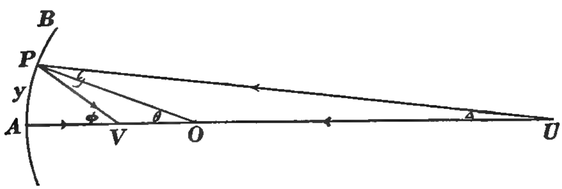


FIG. 241.

highly polished. The *center of curvature O* of the mirror is the center of the spherical shell of which the mirror is a part.

[1] *For the measurement of angles by the sextant, see Manual, Exercise 11.*

The *vertex* A of the mirror is the center of figure of the mirror. A right line AO joining the vertex and the center of curvature is called the *principal axis* of the mirror. From geometry it is evident that any line from O will cut the mirror normally, or a ray of light passing through the center of curvature will strike the mirror at zero incidence and be returned along the same path. Also any ray striking the mirror at any point will form equal angles of incidence and reflection with the radius of the mirror at that point.

Let a luminous point be placed at U (Fig. 241) upon the principal axis UA. Consider two rays UA and UP. Let the first pass through the center of curvature O. After reflection it will return upon its path. Let the second be reflected at any point P, making equal angles with the normal OP, and crossing the first ray at V.

It is required to express the relation between the distances AU and AV, and the radius of the mirror. Let the acute angles at U, O and V be denoted by Δ, θ and ϕ, respectively, and let i denote the angle of incidence and the angle of reflection at P.

Then from the triangles OPU and VPO we have

$$\theta = \Delta + i \tag{462}$$

$$\phi = \theta + i \tag{463}$$

whence

$$\phi - \theta = \theta - \Delta \tag{464}$$

or

$$\Delta + \phi = 2\,\theta \tag{465}$$

Let AU be denoted by p, AV by p', and AO by r. If P be taken very near the point A so that the angles ϕ, θ and Δ are very small, and hence may be set equal to their tangents, we have

$$\frac{AP}{p} + \frac{AP}{p'} = \frac{2\,AP}{r} \tag{466}$$

or

$$\frac{1}{p} + \frac{1}{p'} = \frac{2}{r} \tag{467}$$

a fundamental formula. Since only p, p' and r appear in the

formula, all rays from U will meet after reflection at V, or conversely, all rays from V will meet after reflection at U. Two points so related are called *conjugate points* or *conjugate foci*. The point V is a *real focus*, since the rays actually pass through it after reflection.

In the convex mirror the image is virtual and appears to be behind the mirror. By considering the quantities p' and r negative the above formula may be shown to hold for the convex mirror as well.

444. Discussion of Formula. The formula given in (467) is based upon the assumption that the angular opening of the luminous pencils shall be *small*, that is, that the point P shall lie near the principal axis of the mirror. Within these limits it is true for any value of p and hence for p equal to infinity. In this case

$$p' = \frac{r}{2} \tag{468}$$

This means that for an infinitely distant source of light, *i.e.* for parallel rays, the *focus lies halfway between the center of curvature and the mirror.* This focus for rays parallel to the principal axis is called the *principal focus.*

Again, since the sum of the reciprocals of p and p' is a constant, we see that as p *increases*, p' *decreases*, and *vice versa.* This means that *image and object move at all times in opposite directions.* Suppose p to decrease from plus infinity to r, that is, suppose the object to move from infinity up to the point O, where p is equal to r. We find from the formula that p' is also equal to r. Or *image and object meet at the center of curvature.* As p continues to decrease, the image moves away toward infinity, or p' becomes infinite when p becomes equal to $r/2$. This means that for p equal to $r/2$ the image vanishes. Also for values of p less than $r/2$, the image is *virtual and lies behind the mirror.* In other words, as the object passes the principal focus moving toward the mirror, the image shifts from plus infinity to minus infinity and moves up toward the mirror from behind, and *object and image meet again at the mirror.*

445. Construction of Images in a Concave Mirror. Having given a concave mirror *MN* (Fig. 242), with center of curvature at *C*, and principal focus at *F*, it is required to construct the image of an object *AB*, placed beyond the center of curvature. Since the intersection of two lines is sufficient to locate a point, it will suffice to trace the path of *two rays* from any point on the object in order to locate the corresponding point on the image. Obviously *any two* rays may be chosen for the purpose, making the angles of incidence and reflection equal by construction. It is convenient, however, so to choose the rays from each point that one ray shall lie parallel to the principal axis, and the other shall pass through the center of curvature, since in these two cases the direction of the rays after reflection is perfectly determined, as the first ray will pass through the principal focus and the second will retrace its path.

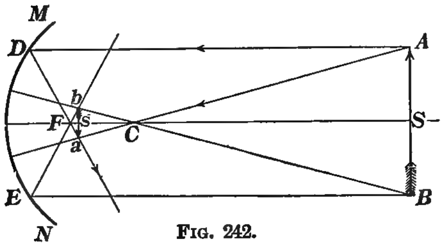


FIG. 242.

Constructing two rays in this manner from the point *A*, it is found that they cross *after reflection* at the point *a*, which is, therefore, the real image of *A*. In the same way the image of *B* is found to lie at *b*, and the image of the object may be sketched in between these two points.

The image seen in a concave mirror, of an object placed ***beyond the center of curvature***, is found to be ***real, inverted, smaller than the object, and located between the center of curvature and the principal focus.*** Also, since rays from *a* and *b* would, after reflection, meet at *A* and *B* respectively, it is clear that *ab* may be considered as an object and *AB* as the corresponding image.

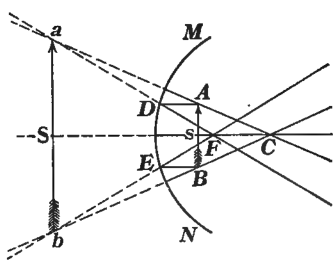


FIG. 243.

In case the object lie between the mirror (Fig. 243) and the principal focus, the reflected rays are *divergent*, and will not meet in front of the mirror. If, however, the reflected rays be produced *backward*, they seem to meet in the virtual points a and b, and the image may be sketched in as shown in the figure. In this case it is seen that the image of an object placed between the principal focus and the mirror is *virtual, erect, larger than the object, and situated behind the mirror.*

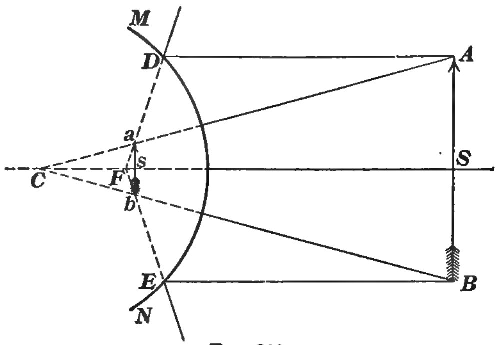


FIG. 244.

In the case of convex mirrors (Fig. 244), the center of curvature and the principal focus are both behind the mirror. Otherwise the construction is the same as in the concave mirror.

Again, if we denote by S and s the points of intersection of the principal axis with the object AB and the image ab (Figs. 242, 243, 244), we have a pair of similar triangles, ACB and aCb, in each figure, from which we may write down at once the proportion

$$\frac{ab}{AB} = \frac{Cs}{CS} \tag{469}$$

which shows that in all cases *the size of the image is to the size of the object directly as their distances from the center of curvature.* They are also proportional to their distances from the vertex of the mirror.

CHAPTER LI

PHENOMENA OF REFRACTION

446. Refraction. When a ray of light passes obliquely from one medium into another of different optical property, it undergoes a change in direction at the surface of separation of the two media. The portion entering the second medium is said to be *refracted.*

If MN (Fig. 245) represent the surface of separation between the two media, where the lower medium is the one of greater optical density, then BA is the *incident* and AC the *refracted* ray. The plane BAD containing the incident ray and the normal AD, at the point of incidence, is the plane of incidence. The plane CAE is the plane of refraction. The angle DAB is the angle of *incidence,* CAE the angle of *refraction,* and HAC the angle of *deviation.*

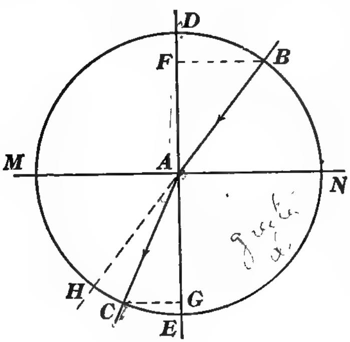


FIG. 245.

The law of refraction, first discovered by Snell in 1621, states that *the refracted ray lies in the plane of incidence and that for light of a definite color, the ratio between the sine of the angle of incidence and the sine of the angle of refraction is a constant, depending only upon the nature of the two media.*

Mathematically the law may be stated

$$\frac{\sin i}{\sin r} = \mu = constant \qquad (470)$$

where i denotes the angle of incidence, and r the angle of refraction.

The quantity μ is termed the *index of refraction* for the two media in question. In case the light passes from vacuum into either medium the ratio is called the *absolute index* of refraction for the medium in question. When it passes from one medium to another it is termed the *relative index*. Again, if the ray BA (Fig. 245) be conceived as passing from air to water, then

$$\frac{\sin DAB}{\sin CAE} = \mu_{aw} \tag{471}$$

where μ_{aw} denotes the relative index of refraction *from air to water*.

If, however, the ray be reversed, it will retrace its path from C to B and we have

$$\frac{\sin CAE}{\sin DAB} = \mu_{wa} = \frac{1}{\mu_{aw}} \tag{472}$$

or the *relative index from water to air is the reciprocal of the index from air to water*. The value of μ_{aw} is 4/3, while μ_{wa} is 3/4. Usually the index is given for the light passing from the rarer to the denser medium and in this case its value is always greater than unity.

***447. Refraction through Plane Parallel Plates.** If a ray of light be successively refracted at the surfaces of one or more plane parallel plates or layers of media, of different refractive indices, several interesting results are obtained. (*a*) *A single plate.* Thus let a ray be refracted through a single plate of glass

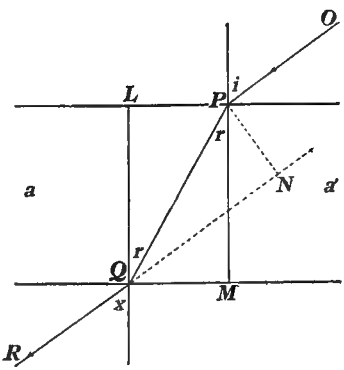


FIG. 246 *a*.

aa' (Fig. 246 *a*), having plane parallel sides. The path of the light is $OPQR$. In this case we have for the refraction at the points P and Q the equations

$$\left.\begin{aligned}\frac{\sin i}{\sin r} &= \mu_{ag} \\ \text{and} \qquad \frac{\sin r}{\sin x} &= \frac{1}{\mu_{ag}}\end{aligned}\right\} \tag{473}$$

by equation (472), whence by multiplication we have

$$\frac{\sin i}{\sin r} \cdot \frac{\sin r}{\sin x} = 1 \tag{474}$$

or

$$\sin i = \sin x$$

and

$$x = i$$

since neither angle is greater than 90°.

From this it is clear that *a ray of light refracted through a medium bounded by plane parallel sides suffers no change in direction, but does undergo a lateral displacement.*

The value of this displacement PN (Fig. 246 *a*) is readily found in terms of the thickness of the plate t, or PM, and the angles i and r. Thus from the triangle PQN we see that

$$PN = PQ \sin PQN = PQ \sin (i - r)$$

further, from triangle PQM, we have

$$PQ = \frac{PM}{\cos r} = \frac{t}{\cos r}$$

whence

$$PN = \frac{t \sin (i - r)}{\cos r} \tag{475}$$

from which we see that the displacement varies directly as the thickness of the plate and also that the displacement increases as the angle of incidence increases.

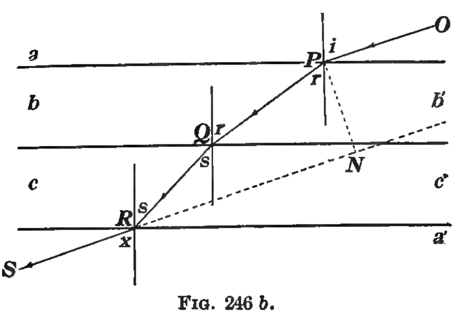


FIG. 246 *b*.

(*b*) *Two plates.* Next let a ray traverse two layers of transparent media, having plane parallel boundaries but different refractive indices. Suppose the ray to enter from the air at

P (Fig. 246 b), and to emerge into the air again at R. To fix our ideas, suppose the layer bb' to be water and the layer cc' to be glass. The path of the ray is $OPQRS$, and the equations for the refraction at P, Q, and R may be written down at once as

$$\frac{\sin i}{\sin r} = \mu_{aw}$$

$$\frac{\sin r}{\sin s} = \mu_{wg}$$

$$\frac{\sin s}{\sin x} = \mu_{ga}$$

Now by experiment we find x is equal to i, whence by multiplication we have

$$\frac{\sin i}{\sin r} \cdot \frac{\sin r}{\sin s} \cdot \frac{\sin s}{\sin x} = 1 = \mu_{aw} \cdot \mu_{wg} \cdot \mu_{ga}$$

or

$$\mu_{wg} = \frac{1}{\mu_{aw} \cdot \mu_{ga}} = \frac{\mu_{wa}}{\mu_{ga}} = \frac{\mu_{ag}}{\mu_{aw}} \qquad (476)$$

From this we see that *the relative index of refraction* μ_{wg}, *for any two media as from water to glass, may be expressed as the ratio between the relative indices of the same two media to some third medium,* $\frac{\mu_{wa}}{\mu_{ga}}$, *as air; or as the inverse ratio between the relative indices of a third medium (air) to the two media in question,* $\frac{\mu_{ag}}{\mu_{aw}}$. On substituting numerical values for $\mu_{ag} = 3/2$, and for $\mu_{aw} = 4/3$, we have for the relative index of refraction from water to glass

$$\mu_{wg} = \frac{\frac{3}{2}}{\frac{4}{3}} = \frac{9}{8} = 1.125$$

Further, if for the third medium vacuum be chosen, then the relative indices μ_{ag} and μ_{aw} become the *absolute indices* for glass and water respectively (written simply μ_g and μ_w), and the equation for refraction from water to glass may be rewritten in the form

$$\mu_{wg} = \frac{\sin r}{\sin s} = \frac{\mu_g}{\mu_w}$$

or

$$\mu_w \sin r = \mu_g \sin s \qquad (477)$$

or, in the case of refraction from medium to medium, *the product of the absolute index of a medium and the sine of the angle in that medium remains constant from medium to medium.*

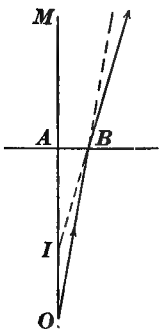


FIG. 247.

448. Refraction at a Plane Surface. Consider a luminous point O (Fig. 247) in the denser medium. A ray emerging along the line OB will be refracted on entering the rarer medium as shown in the figure. Draw MO normal to the surface. Then the angle BOA is equal to i, and BIA is equal to r, in the rarer medium, consequently

$$\frac{\sin BOA}{\sin BIA}=\frac{1}{\mu}$$

since the refraction is from dense to rare. Substituting from the right triangles BAI and BAO the values of the sines of the angles, we have

$$\frac{\dfrac{AB}{BO}}{\dfrac{AB}{BI}}=\frac{1}{\mu} \tag{478}$$

whence

$$\frac{1}{\mu}=\frac{BI}{BO} \tag{479}$$

or

$$\mu=\frac{BO}{BI} \tag{480}$$

If the point B approach very near to A, we may put BO equal to AO, and BI equal to AI, whence

$$\mu=\frac{AO}{AI} \tag{481}$$

If the media be air and water, then a point situated at O would appear to be at I, to an observer directly over it. But μ from air to water is $\frac{4}{3}$, hence AI is $\frac{3}{4}$ AO, or the object appears to be but $\frac{3}{4}$ its real distance below the surface. Conversely, to an eye under water, the point M would appear to be $\frac{4}{3}$ its real distance away from the surface.[1]

[1] *For method of determining μ by this principle, see Manual, Exercise 87.*

449. Critical Angle. When light passes from a denser to a rarer medium, as from water to air, it is refracted away from the normal, the angle of refraction being greater than the angle of incidence. As the angle of incidence increases, the angle of refraction increases more rapidly, until for a certain limiting angle of incidence in the denser medium, the angle of refraction becomes 90°, or the refracted ray OB *grazes the surface of separation between the two media.* Thus (Fig. 248), the ray RO is refracted along the line OS, while for the ray LO the refracted ray becomes OB. Obviously light incident at any angle greater than LON' cannot emerge from the water, but will be reflected back again. Thus a ray meeting the surface along the line IO will be reflected to I', and since no light can emerge from the water at this incidence, it suffers *total internal reflection.*

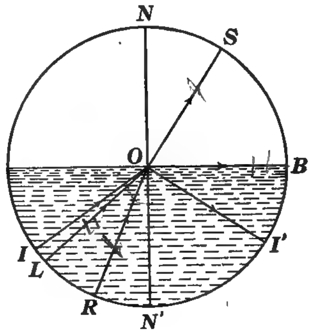


FIG. 248.

The angle LON' is called the *critical angle*, since for this angle of incidence refraction ceases. The critical angle may be defined as *that angle of incidence in the denser medium which corresponds to an angle of refraction of* 90°. If α be the critical angle, and μ the index of refraction for the two media, then,

$$\frac{\sin \alpha}{\sin 90^\circ} = \frac{1}{\mu} \tag{482}$$

or

$$\sin \alpha = \frac{1}{\mu} \tag{483}$$

That is, *the sine of the critical angle is the reciprocal of the index of refraction.*

The critical angle for water is $\sin^{-1}(\frac{3}{4}) = 48^\circ\ 27'$. This means that a diver on the bottom of the ocean can see out through a cone whose base rests in the surface of the water and whose semi-aperture is about $48^\circ\ 27'$. If he look toward the surface at an angle greater than this, he can see only the reflection

of objects upon the bottom. Not only this, but since all the light that can reach him must come in through this same cone, he will see all the stars in the heavens crowded into a circle whose center is the zenith, and whose radius subtends an angle of about 48° 27′.

Familiar examples of total reflection are seen in the case of a tumbler of water held above the head; the under side of the upper surface of the water gives a clear, mirror-like image of the objects below it; — also in the case of a bubble of air in water, or a test tube containing air immersed in water, which when viewed at a certain angle will seem to be filled with quicksilver.

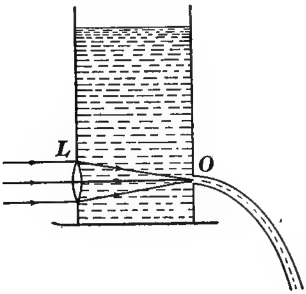


Fig. 249.

A beautiful illustration of total reflection is furnished by a stream of water flowing from a tank (Fig. 249), having in the side opposite the orifice a lens *L*, by means of which a strong beam of light may be concentrated within the jet. The light strikes the surface of the jet at an angle greater than the critical angle and is held prisoner within the stream, until, after repeated reflection, it strikes the bottom of the receptacle, where it shows as a bright spot. A goblet held in the stream is filled with light. The electrical illumination of fountains depends upon this principle.

CHAPTER LII

PRISMS AND LENSES

450. Refraction through a Prism. A prism is a transparent substance bounded on two sides by plane surfaces meeting at an angle. The line of intersection of these two planes is called the edge of the prism, and the face opposite this edge is called the base. Light on entering the prism at I (Fig. 250) is refracted *toward* the normal, along the line IE. On emerging at E the ray is refracted *from* the normal, along EO. The ray is *thus bent toward the base of the prism in each case.* Let i and r denote the angles of incidence and refraction at the first face of the prism, and r' and i' the angles of incidence and emergence at the second face. Then, since the angle between the faces of the prism is the same as that between the normals NI and ME, it follows that

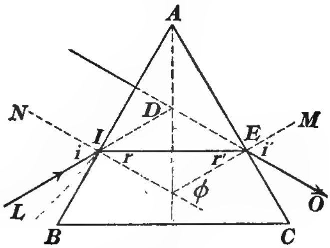


FIG. 250.

$$A = \phi = r + r' \tag{484}$$

Also the deviation at the first face is $i - r$, and at the second face $i' - r'$, or the total deviation D is

$$D = i - r + i' - r' = i + i' - A \tag{485}$$

It may be shown both from theory and by experiment that for every prism there is a minimum value of the deviation, below which it cannot fall, and that this deviation is a minimum when the path of the ray through the prism is symmetrical, or when

$$i = i' \text{ and } r = r' \tag{486}$$

Upon this condition we have

$$2r = A$$

or
$$r = \frac{A}{2} \tag{487}$$

and
$$2i = A + D$$

or
$$i = \frac{(A + D)}{2} \tag{488}$$

Substituting,
$$\mu = \frac{\sin i}{\sin r} = \frac{\sin \frac{1}{2}(A + D)}{\sin \frac{1}{2} A} \tag{489}$$

This formula is used in determining the indices of refraction of transparent solids in the form of prisms, and of liquids inclosed in a hollow prism with glass sides. The angle D denotes the angle of minimum deviation for the color under investigation.[1]

451. Prisms of Large and Small Angle. In prisms of very small angle, we may substitute in the foregoing formula, instead of the sines of the angles, the angles themselves, and the approximate formula becomes

$$\mu = \frac{A + D}{A} \tag{490}$$

and
$$D = A(\mu - 1) \tag{491}$$

If a perpendicular be dropped from A (Fig. 250) upon IE, when $r = r'$, then each half of the angle A is equal to the angle of refraction r. Now, since r must in all cases be less than the critical angle for the substance of the prism, it is clear that the angle A *cannot exceed twice the critical angle if the light is to emerge from the prism.*

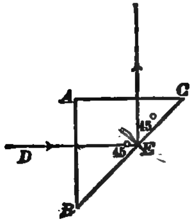


Fig. 251.

In the case of crown glass the critical angle is about 41° 16′. If a prism of crown glass be cut in the form shown in Fig. 251, the angle A will exceed twice the critical angle, and a ray DE, incident normally upon the face AB, will be totally reflected at E, and emerge normally from the

[1] *For method of determining μ in case of a prism, see Manual, Exercise 89.*

face AC. Such prisms are frequently employed in optical apparatus.

452. The Abbé-Littrow Principle. If the prism ABC (Fig. 250) were split along the perpendicular from A upon IE, and the new surface were polished and silvered, it is easily seen (Fig. 252) that the refracted ray would strike the silvered surface normally, retrace its path, and return along its original direction. The ray would have traversed the half prism twice and hence would have undergone the same changes as though it had emerged from the opposite side.

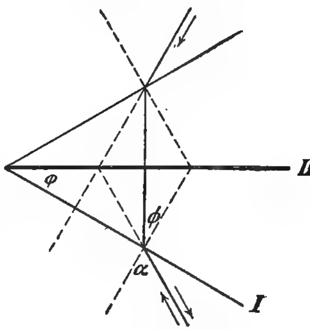


FIG. 252.

In the Abbé spectroscope, light from a slit in the focal plane of a telescope, emerging from the objective as parallel rays, is allowed to fall upon such a half prism, and the returning rays are received again into the telescope, where the refracted image of the slit is made to coincide with the slit itself. By this ingenious device the instrument is much simplified and the labor of making measurements is greatly reduced, since, when the refracted image of the slit coincides with the slit, *the position of minimum deviation is attained.*

In this case we have

$$\mu = \frac{\sin \alpha}{\sin \phi} \tag{492}$$

453. Refraction through a Thin Lens. A lens is a portion of a refracting medium bounded by two surfaces of revolution which have a common axis called the *axis of the lens*. In the majority of cases the lens is bounded by spherical surfaces, although in large lenses the spherical surfaces must be changed or "corrected" in order to overcome certain optical defects which will be noticed later. By a *thin lens* is meant one whose thickness is small in comparison to its radii of curvature.

In Fig. 253, let O be the center of curvature of the spherical surface BB', forming the front surface of a thin lens. Let U be

a point source of monochromatic light, of index μ. Let one ray meet the lens at A and a second ray at P. After refraction the ray UP will have the direction VP, while the ray UA, passing through O, strikes the surface normally and will enter the denser medium without change of direction. Consequently to an eye situated *in the glass* these two rays would seem to proceed from the point V. The points U and V are then conjugate foci *for the first surface.* From P draw the radius through O and produce it to meet perpendiculars let fall from V and U. Then $i = UPO$, and $r = VPO$, and by the law of refraction we have

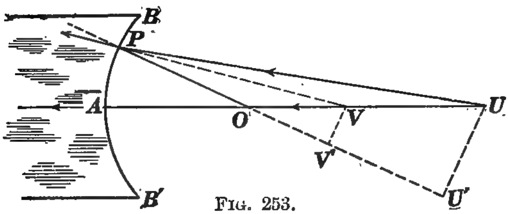


FIG. 253.

$$\mu = \frac{\sin i}{\sin r} = \frac{\sin UPO}{\sin VPO} = \frac{UU'}{PU} \div \frac{VV'}{PV} \tag{493}$$

or

$$\mu = \frac{UU'}{VV'} \cdot \frac{PV}{PU}$$

But from the similar triangles $VV'O$ and $UU'O$, we have

$$\frac{UU'}{VV'} = \frac{OU}{OV} \tag{494}$$

whence

$$\mu = \frac{OU}{OV} \cdot \frac{PV}{PU} = \frac{OU}{OV} \cdot \frac{AV}{AU} \tag{495}$$

if P be taken very near to the point A.

If now we set AU equal to p, AV equal to p', and AO equal to r_1, *the radius of the surface first struck by the light*, we have

$$\mu = \frac{p - r_1}{p' - r_1} \cdot \frac{p'}{p} \tag{496}$$

whence, clearing of fractions, dividing both sides by $pp'r_1$, and transposing, we have

$$\frac{\mu}{p'} - \frac{1}{p} = \frac{\mu - 1}{r_1} \tag{497}$$

If for μ we substitute -1, this formula reduces to that for the concave spherical mirror, since for reflection, μ must be unity, and the negative sign indicates a reversal of the direction of the ray.

If now the refracted ray meet the second surface of the lens, whose radius is r_2, and whose center of curvature lies on OA, it will be refracted a second time, passing now *from a denser to a rarer medium.* The new index of refraction will therefore be $1/\mu$. The point image V now becomes the object, and p' becomes the object distance. Let q be the distance of the final image from the second surface and *neglect the thickness of the lens.*

Then
$$\frac{\frac{1}{\mu}}{q}-\frac{1}{p'}=\frac{\frac{1}{\mu}-1}{r_2} \tag{498}$$

or
$$\frac{1}{q}-\frac{\mu}{p'}=\frac{1-\mu}{r_2} \tag{499}$$

Adding (497) and (499), we have
$$\frac{1}{q}-\frac{1}{p}=(\mu-1)\left(\frac{1}{r_1}-\frac{1}{r_2}\right) \tag{500}$$

This formula is to be understood as an approximation, giving accurate values only in the case of pencils of small angular opening, and for thin lenses.

454. Discussion of Formula. If the source be at an infinite distance, formula (500) of the last article becomes
$$\frac{1}{q}=(\mu-1)\left(\frac{1}{r_1}-\frac{1}{r_2}\right) \tag{501}$$

If we set this value of q equal to f, we may write
$$\frac{1}{f}=(\mu-1)\left(\frac{1}{r_1}-\frac{1}{r_2}\right) \tag{502}$$

where f is called the *focal length* of the lens, *i.e.* the distance from the rear surface of the lens to the focus of rays parallel to

the *principal axis.* Such a focus is called the *principal focus*, and its conjugate point lies at infinity.

In the derivation of the foregoing formulae it has been assumed that in all cases the light proceeds *from right to left*, further that all distances are to be measured *from* the surfaces of the lens *to* the points in question. Accordingly, distances measured to the right are *positive* and those to the left are *negative.* A more general statement may be made as follows:

Distances measured toward the source of light are to be taken as positive and those measured away from the source as negative; the measurements in all cases to be taken from the optical surface to the point in question.

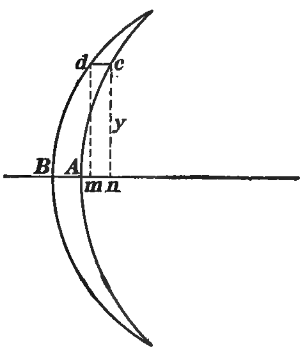


FIG. 254.

455. The Sign of the Quantity f. It is proposed to determine when the focal length of a lens is to be considered *positive* and when *negative.* Consider the meniscus-shaped lens BA (Fig. 254). Set dc equal to T, the thickness of the lens, measured parallel to the axis at a distance cn equal to y, from the axis. Let t equal BA, the thickness of the lens at the axis; *i.e.* the value of T for y equal to zero.

Let r_1 and r_2 be the radii of the first and second surfaces respectively. Then, by geometry, $cn^2 = An\,(2\,r_1 - An)$ or, neglecting An^2 we have

$$An = \frac{y^2}{2\,r_1} \tag{503}$$

and

$$Bm = \frac{y^2}{2\,r_2}$$

Also

$$Bm + T = An + t \tag{504}$$

or

$$T - t = An - Bm = \frac{y^2}{2}\left(\frac{1}{r_1} - \frac{1}{r_2}\right) \tag{505}$$

whence by equation (502)

$$T - t = \frac{y^2}{2(\mu - 1)} \cdot \frac{1}{f} \tag{506}$$

Hence, since μ is greater than unity, the sign of f is positive or negative according as T is $>$ or $<$ t; that is, according as the lens is *thicker at the edges than at the center* or *vice versa.*

A lens whose focal length is *positive*, that is, a lens thicker at the edges than at the center, and whose principal focus lies on the same side of the lens as the source of light, is called a *concave* or *diverging* lens. A lens thicker at the *center* than at the edges has its focal length *negative*, and is called a *convex* or *converging* lens. Examples of the various forms of convex and concave lenses are shown in Fig. 255.

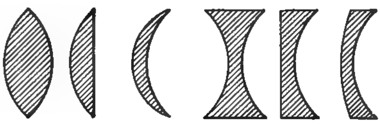


FIG. 255.

In case the two radii r_1 and r_2 are equal, the quantity in the parenthesis in equation (505) can never reduce to zero, *since the radii are always measured in opposite directions.* For a *double convex lens*, the expression

$$\frac{1}{f} = (\mu - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$$

reduces to

$$-\frac{1}{f} = (\mu - 1)\frac{2}{r_1} \tag{507}$$

From this it is evident that r_1 is *negative*, as it should be, since *the center of the surface first struck by the light lies to the left or on the opposite side of the lens from the source of light.*

In the case of a double concave lens we have

$$\frac{1}{f} = (\mu - 1)\frac{2}{r_1} \tag{508}$$

or r_1 is *positive* in this case, and *the center of the first surface lies on the same side as the source.*

456. Discussion of Lens Formula. Concave Lenses. Since in the expression

$$\frac{1}{f} = (\mu - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$$

the right-hand member contains only constant quantities for any given lens, and for light of a definite color, it follows that the focal length of a lens is a constant, and depends simply upon the material of the lens and upon the curvature of its surfaces. Again, since

$$(\mu - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) \text{ is equal to both } \frac{1}{f} \text{ and to } \frac{1}{q} - \frac{1}{p}$$

we may write

$$\frac{1}{q} - \frac{1}{p} = \frac{1}{f} \tag{509}$$

It has already been shown that for a concave lens f is always positive. Now since p, the object distance, is essentially positive, it follows that *q must be positive and less than p.* This means that the image lies on the same side of the lens as the object, that it is virtual and nearer the lens than the object. This is shown in Fig. 256, where the concave lens MN produces a *virtual, diminished, erect* image ab, of the object AB, and the distances OK, Ok and OF represent the quantities p, q and f.

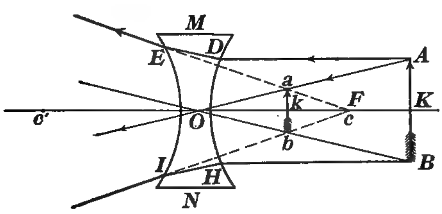


FIG. 256.

The foregoing figure is easily constructed by taking from each point two rays, whose paths after refraction are definitely determined. These rays are (a) a ray parallel to the principal axis, which, after refraction, either passes *through*, or seems to proceed *from*, the principal focus; (b) a ray which passes through the optical center O, of the lens. All rays passing through this point suffer no deviation. In the case of double concave or double convex lenses this point lies at the center of the lens. In lenses of other forms the location

of this point is determined from formulae too complex for an elementary text.

457. Discussion of Lens Formula. Convex Lenses. In convex lenses it has been shown that f is negative and the formula becomes

$$\frac{1}{q} - \frac{1}{p} = -\frac{1}{f} \tag{510}$$

Three cases require consideration:

(*a*) *When* $p < f$, or $1/p > 1/f$, q must be *positive* and *greater* than p, and *image and object lie on the same side of the lens.* The image is then virtual, erect and magnified as shown in Fig. 257. In this case if we call the intersections of image and object with the axis of the lens, k and K, we may set OK equal to p, Ok equal to q, and OF equal to f. An example of this case is found in a single lens used as a simple magnifier.

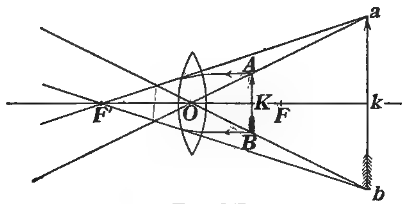


FIG. 257

(*b*) *When* $p > f$, or $1/p < 1/f$, q must be negative, since by (510) $1/q - 1/p$ is equal to a negative quantity. This means that the image lies on the opposite side of the lens from the object, that it is real, inverted and smaller or larger than the object according as p is greater or less than q. This case is illustrated in Fig. 258.

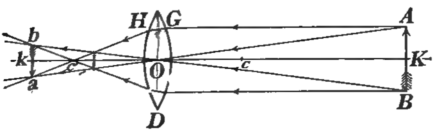


FIG. 258.

Since, for a real image in the case of a convex lens, q is negative, we may rewrite the formula after changing signs,

$$\frac{1}{q} + \frac{1}{p} = \frac{1}{f} \tag{511}$$

(*c*) *When* $p = q$, we have

$$\frac{2}{p} = \frac{1}{f}, \text{ or } p = 2f \tag{512}$$

This means that the image is at the same distance behind the lens as the object is in front of it, that it is inverted, real and the same size as the object. This case is applied in a method for measuring the focal length of a convex lens. The lens, an object, and a screen are mounted upon an optical bench and the screen moved until the image is the same size as the object. The distance from object to screen is then four times the focal length of the lens.

It remains to be noted that in all cases of images formed by lenses, *the size of the image is to the size of the object directly as their distances from the center of the lens.*

458. Image and Object at a Fixed Distance. It appeared in the previous article that the real image formed by a convex lens may be larger or smaller than the object. If a convex lens, an object and a screen be mounted upon an optical bench and the distance between object and screen be made more than four times the focal length of the lens, the two images may be thrown upon the screen in succession by simply moving the lens. In the first case the lens is nearer the object and the image is correspondingly magnified. In the second case the two distances p and q are interchanged and the lens is nearer the image. If we call the distance from the object to screen l, and the difference in the two settings of the lens a, then

$$\left.\begin{aligned} q+p&=l \\ q-p&=a \end{aligned}\right\} \qquad (513)$$

whence $$q=\frac{l+a}{2} \quad \text{and} \quad p=\frac{l-a}{2}$$

Substituting in $$\frac{1}{q}+\frac{1}{p}=\frac{1}{f}$$

we have $$\frac{1}{f}=\frac{1}{\frac{l+a}{2}}+\frac{1}{\frac{l-a}{2}}$$

or $$f=\frac{l^2-a^2}{4l} \qquad (514)$$

This is a more accurate method of measuring f than that given in the previous article, since it is difficult to focus accurately in the case where the two positions of the lens coincide, *i.e.* where $l = 4f$.[1]

459. Constants of Thick Lenses. In the case of lenses whose thickness cannot be neglected, recourse is had to the method of Gauss, who first demonstrated that the formula

$$\frac{1}{q} - \frac{1}{p} = \frac{1}{f}$$

may be applied to thick lenses through the introduction of certain cardinal points to be determined for each lens. The constants of a thick lens are defined as follows:[2]

Conjugate planes are planes normal to the axis of a system, so related to each other that to every point in the plane in object space there corresponds a point in the plane in image space.

Magnification. If y and y' be the respective distances of a point object and its point image from the axis of the system, then the ratio y'/y is called the *magnification, m.*

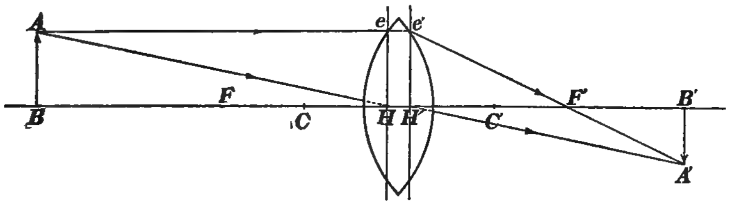


FIG. 259.

The principal planes eH, e'H' (Fig. 259) are the two conjugate planes in every system, in which object and image are of the *same size* and *stand in the same position.* Since in the

[1] *For method of determining focal lengths of lenses, see Manual, Exercise 82.*

[2] For a detailed discussion of this subject the student is referred to Müller-Pouillet's *Lehrbuch der Physik,* 9th ed., vol. 2, part I, pp. 85–180, or Drude's *Theory of Optics,* pp. 17–30.

principal planes the magnification

$$m = \frac{y'}{y'} = +1 \tag{515}$$

the principal planes are sometimes called *unit planes* or planes of *unit magnification.* The points of intersection H and H' of the principal planes with the axis of the system are called the *principal points*, or *unit points.*

The *focal points* F and F' are the points on the axis in which all rays parallel to the axis in image and object space respectively meet after passing through the system.

The *focal planes* are the two planes normal to the axis at the focal points, and having their conjugates at infinity.

The *focal lengths* f and f' of a system for object and image space, respectively, may be defined as the distances from the focal points F and F', to the principal points H and H'.

If in Fig. 259 we set HB equal to p, $H'B'$ equal to q, and $H'F'$ equal to f, then our previous formula

$$\frac{1}{q} - \frac{1}{p} = \frac{1}{f}$$

for convex lenses still holds. This means that the distances from object to lens and from lens to image are now to be measured from the principal points instead of from the surface of the lens as heretofore.

In the case of crown glass lenses of equal curvature the principal points are located within the lens, at one third the thickness of the lens from the curved surfaces. In the case of plano-concave or plano convex lenses of crown glass, one of the principal points lies on the curved surface and the other at one third the thickness of the lens from the curved surface. In thin lenses the two principal points coincide at the center of the lens. In lenses of the meniscus form one or both principal points may lie outside the lens.

***460. Geometrical Significance of Focal Lengths.** In any lens system (Fig. 260), let F and F'', H and H', represent the focal

points and principal points, in object space and image space respectively. Let PFA, a ray through F, the focal point in object space, make an angle u with the axis, and let $A'P'$ be the corresponding or conjugate ray in image space. In general, parallel rays in object space must intersect in some point in the focal plane through F' in image space. Let this point be distant y'

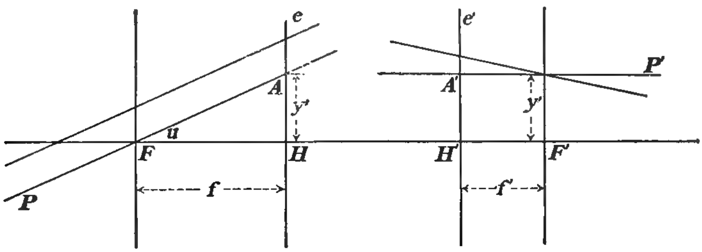


FIG. 260.

from the axis. The value of y' evidently depends upon the angle of inclination u which the incident ray makes with the axis. If u be zero, then y' is zero, that is, rays parallel to the axis intersect in F'.

But in the case of ray PFA which passes through F, the focal point in object space, and cuts the principal plane H in A, u is not zero. Its conjugate ray $A'P'$ must evidently be parallel to the axis, and, from the definition of principal planes, A and A' are equidistant from the axis.

Hence y', the distance from the axis to the image formed by a parallel beam incident at an angle u, is shown from the figure to be

$$y' = f \tan u, \tag{516}$$

and by symmetry we may write

$$y = f' \tan u', \tag{517}$$

where u' is the angle under which a ray parallel to the axis in object space cuts the axis in image space.

***461. Gauss's Definition of Focal Lengths.** Let P (Fig. 261) represent the position of an infinitely distant object, *e.g.* the

sun, from the system S. The rays proceeding from all points of the sun may be regarded as parallel. The system S receives therefore, simply cylinders of rays which are focused in the corresponding points of the rear focal plane through F'. This produces in F' a small image of the sun. Let h' represent its diameter, and let u be the angle which the extreme rays from the sun form with the axis. Then the *front focal length*,

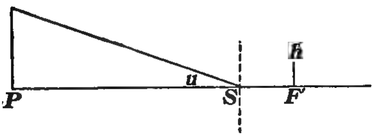


FIG. 261.

$$f=\frac{h'}{\tan u} \tag{518}$$

is equal to the diameter of the sun's image divided by the tangent of the angle subtended by the sun's disk. This latter angle is termed the *visual angle* or the *apparent magnitude.*

The focal length may therefore be defined as follows :

The focal length of object space is the quotient obtained by dividing the linear magnitude of the image of an infinitely distant object by the tangent of the angle subtended by that object, or

$$f=\frac{y'}{\tan u} \tag{519}$$

For the rear focal length we may write

$$f'=\frac{y}{\tan u'} \tag{520}$$

or *th focal length of image space is equal to the distance between the axis and any ray parallel to it in object space, divided by the tangent of the inclination of its conjugate ray.*

***462. Determination of Focal Lengths.** Let LL' and QQ' (Fig. 262) be two pairs of conjugate planes with reference to a system whose focal points are F and F'.

Let the two rays I and II intersect at P and their conjugates I' and II' at P'. Ray I is parallel to the axis in object space and passes through F' at an angle u'. Ray II passes through

F at an angle u, and is parallel to the axis in image space. The planes L and L' may be considered as the first pair and Q and Q' as the second pair of conjugates, to which the subscripts of x and y correspond.

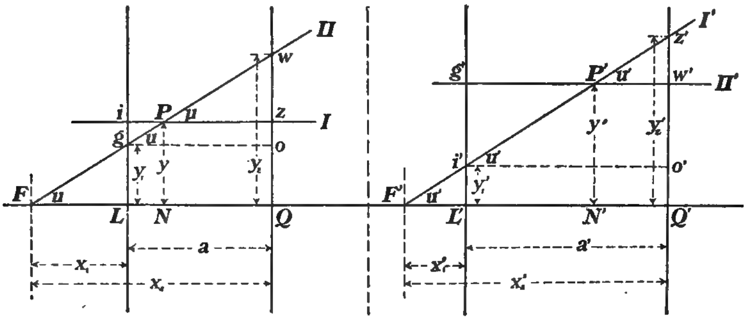


FIG. 262.

Through a special choice of rays we have for ray I in object space, $iL = zQ = PN$, or $y_1 = y_2 = y$, and consequently the magnification may be expressed by

$$\frac{y'_1}{y} = m_1, \text{ and } \frac{y'_2}{y} = m_2$$

In a similar manner we have for ray II in image space $g'L' = w'Q' = P'N'$, or $y'_1 = y'_2 = y'$, and

$$\frac{y_1}{y'} = \frac{1}{m_1}, \ \frac{y_2}{y'} = \frac{1}{m_2} \tag{521}$$

From the triangle wgo, we have

$$\frac{ow}{og} = \tan u = \frac{y_2 - y_1}{a} \tag{522}$$

multiplying both numerator and denominator by y', we have

$$\tan u = \frac{y'}{a} \cdot \frac{y_2 - y_1}{y'} \tag{523}$$

but, since
$$\frac{y_2}{y'}=\frac{1}{m_2}, \text{ and } \frac{y_1}{y'}=\frac{1}{m_1}$$

therefore
$$\tan u=\frac{y'}{a}\left(\frac{1}{m_2}-\frac{1}{m_1}\right) \tag{524}$$

or
$$\frac{a}{\frac{1}{m_2}-\frac{1}{m_1}}=\frac{y'}{\tan u}=f=\frac{x_2-x_1}{\frac{1}{m_2}-\frac{1}{m_1}} \tag{525}$$

In a similar manner from triangle $z'i'o'$

$$\tan u'=\frac{y'_2-y'_1}{a'} \tag{526}$$

or
$$\tan u'=\frac{y}{a'}\cdot\left(\frac{y'_2-y'_1}{y}\right)=\frac{y}{a'}(m_2-m_1) \tag{527}$$

whence
$$\frac{a'}{m_2-m_1}=\frac{y}{\tan u'}=f'=\frac{x'_2-x'_1}{m_2-m_1} \tag{528}$$

In practice a lens system is set up and the image of an object of known dimensions is measured for two positions of the object L and Q, whose distance apart, a or x_2-x_1, can be accurately determined. From these data the values of m_1 and m_2 can be computed and inserted in formula (525).

In the case of a microscope objective, the magnification of a known object is determined, first with the draw tube of the microscope pushed in, and next with it drawn out through a known distance a', the positions of the images corresponding to L' and Q' (Fig. 262). The resulting values of m_1 and m_2 are then substituted in formula (528). This principle finds application in the Abbé focometer and in related methods of determining the focal lengths of lens systems.

463. Spherical Aberration. It is to be noted that *the formulae for spherical mirrors as well as that for refraction through a thin lens are accurate only for pencils of small angular aperture.*

In cases where the opening of the mirror or lens is no longer small, the focus for parallel rays is no longer a point. Rays striking the reflecting or refracting surface at a distance from the axis are brought to a focus sooner than those rays lying nearer the axis. This results in confusion and distortion of the image, which is called *spherical aberration.*

In the case of astronomical specula where a large reflecting surface is essential for light gathering purposes, the spherical form cannot be used. The surface of the mirror must be of parabolic section, since in the case of the parabola any ray parallel to the axis passes, after reflection, exactly through the focus. In compound lenses the defect of spherical aberration is avoided by combining two or more spherical surfaces so that their respective aberrations may annul each other. In a simple lens the defect is partially removed by the use of a diaphragm which stops off the rays from the edges of the lens, thus leaving only the central part of the lens effective.

Since the image of a star is always a point on the axis of the lens, the axial aberration is of chief importance in astronomical work. In photographic lenses, however, it is necessary to recognize and correct for *five different kinds of spherical aberration,* and the manufacture of high grade photographic lenses becomes correspondingly complicated. The discussion of the formulae by means of which these corrections are effected is too difficult for an elementary text.

Problems

1. Two plane mirrors are placed parallel and facing each other at a distance of 20 cm. Required the distance of the first three images, in each mirror, of an object placed 8 cm from one of the mirrors.

Ans. 8, 32, and 48 cm from first; 12, 28, and 52 cm from second.

2. A mirror is made to revolve about a vertical axis 25 times a second. If a horizontal beam of light be allowed to fall on the mirror from a fixed source, required the velocity at which the reflected beam would traverse a circle 78 cm in diameter having its center on the axis of the mirror.

Ans. 1.23×10^4 cm/sec.

3. Let AB and CB be two mirrors inclined to each other at an angle x. Also, let p' be the image in AB of any point p, placed between the mirrors,

and p'' the image of p' in CB. Show that the angular separation of pBp'' is twice the angle between the mirrors.

4. An object 0.96 cm long is placed at a point 35 cm in front of a concave mirror having a focal length of 30 cm. Required the size and position of the image. *Ans.* 5.76 cm long; 210 cm in front.

5. What is the radius of a spherical mirror which forms an image at a distance of 46.2 cm in front of the mirror when the object is placed 153 cm from the vertex. *Ans.* $R = 71.0$ cm.

6. Compute the size of the image of the sun formed by a mirror having a radius of 275 cm, the angular diameter of the sun being taken as 32 min. *Ans.* 1.28 cm.

7. If an eye immersed in a fluid whose index of refraction is 1.42, look out through a horizontal surface, what will be the greatest apparent zenith distance of a star? *Ans.* 44° 46′.

8. Find the radius of the circle on the upper surface beyond which light waves, emitted by a luminous point at the bottom of a layer of liquid 4.2 cm deep and having an index of refraction of 1.25, will cease to emerge. *Ans.* 5.6 cm.

9. When a layer of liquid 4.65 cm deep is poured upon a dot in the bottom of a glass cup, the position of its image, as found by the necessary change in the focus of a microscope, is 1.37 cm above the bottom. What is the index of refraction of the liquid? *Ans.* $\mu = 1.41$.

10. What would be the minimum deviation produced by a prism whose angle is 1°.3, for which $\mu = 1.54$? *Ans.* 42′.12.

11. The minimum deviation produced in monochromatic light by a prism whose angle is 45°.05 is 26°.67. What is the index of refraction? *Ans.* $\mu = 1.530$.

12. The radii of curvature of a thin double convex lens are 46.4 cm, and the index of refraction 1.53. What is its focal length? *Ans.* $f = -43.8$ cm.

13. Required the focal length of a thin lens which forms an image at a distance of 30.3 cm behind the lens, when the object is placed 91.1 cm in front. *Ans.* $f = -22.7$ cm.

CHAPTER LIII

DISPERSION

464. Dispersion and Recomposition of Light. When a ray of sunlight is admitted through a narrow vertical slit into a darkened room and passed through a long focus lens, a sharply defined, white image of the slit is projected upon a screen placed near and parallel to the opposite wall of the room. If now a glass prism with its refracting edge parallel to the slit be placed just inside the focus of the lens, the light is refracted from its original direction toward the base of the prism, and is spread out into a wedge-shaped beam having its apex at the prism and its base upon the screen. The screen should now be moved so as to stand normally to the refracted light. The original white image of the slit has now been replaced by *a series of overlapping, colored images*, forming a continuous band of brilliant color, in which the tints vary insensibly from red nearest the first position of the slit image, through all shades of orange, yellow, green, blue and indigo, to violet at the end of the band farthest from the first position.

Such a band of color is called a *spectrum*. If sunlight be used, it is called a *solar spectrum*. The separation of white light into its component colors is called *dispersion*.

The dispersion of light was first explained by Newton in 1666, who announced as the result of his experiments, that "rays of differently colored light have different degrees of refrangibility." He therefore concluded that white light is to be regarded as a mixture of an endless series of tints, among which the so-called "colors of the rainbow," violet, indigo, blue, green, yellow, orange and red, were selected by him as distinctive. He also proved that after light has been dispersed

by passing it through a prism, it suffers no further change in color on passing through a second prism.

Newton further showed that if the various colors of the spectrum be recombined by any means, as by reflecting them from a series of mirrors, or by receiving them upon a concave mirror and focusing them upon a single point, the result was white light as in the original source.

This experiment may be readily performed by allowing the light from the first prism to traverse a *second, similar prism* (Fig. 263), with its refracting edge turned in the opposite direction. If the prisms be exactly similar and the adjustment be correct, the image of the slit falls in its original position and is perfectly white. The effect of the combined prisms is thus seen to be equivalent to that of a plate with parallel sides.

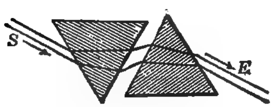


FIG. 263.

In accordance with the physical explanation of refraction to be given later, it may be shown that the varying refrangibility of the different colors is due to the different retardations experienced by these colors in traversing the denser medium. It has been conclusively proven that in a denser medium the speed of violet light is less than that of red light. In its passage through the prism, therefore, the violet light is most retarded and is consequently most refracted, while the red light is least deviated from its original direction. For any two media the index of refraction is, therefore, not a constant, but increases from the red to the violet end of the spectrum.

465. The Fraunhofer Lines. If the slit in the foregoing experiment be made narrower, the overlapping of the spectral images is diminished, since the narrower the slit the smaller is the required difference in refrangibility between two adjacent colors in order to separate their corresponding images in the spectrum. Wollaston, in 1802, was the first to use a narrow slit as a source of light and to secure an approximately *pure spectrum*, that is, a spectrum in which the various colors are almost completely separated from each other. He used a slit

about one twentieth of an inch wide, which he viewed at a distance of 10 or 12 feet, through a prism held near the eye. By thus receiving the spectrum directly into the eye he was able to detect certain dark lines crossing the solar spectrum, which he believed to mark the limits of the spectral colors.

In 1814, Fraunhofer, working independently of Wollaston, rediscovered these dark lines, made a map showing their relative positions, and designated the more prominent lines by letters of the alphabet. Thus (Fig. 264), the *A* line (not shown in figure) lies in the extreme red, the *B* and *C* lines in the medium

FIG. 264.

and lighter red, the *D* line in the yellow, the *E* in the green, *F* in the greenish blue, *G* in the deep blue, while the two *H* lines mark the limit of the violet end of the spectrum for the eyes of most observers.

Fraunhofer also used a single prism at a distance from a narrow slit, but he received the spectrum into a telescope, which was first sharply focused upon the slit before the prism was put in place. By this means he was enabled to observe with such accuracy that he counted 754 lines between *B* and *H*, and located with certainty the positions of 350 of them upon his map.

The Fraunhofer lines in the solar spectrum are of extreme importance in the study of optics, both from a practical and a theoretical point of view. They afford a ready and accurate means of designating lights of definite colors, and are constantly used as reference lines in the determination of refractive indices. On the other hand they indicate the partial absence of certain colors in the light of the sun, and are to be regarded as dark images of the slit. These lines not only offer a starting point for the study of the nature of light, but lead to important conclusions concerning the constitution and condition of the sun itself. It is to be noticed that the number of these dark lines observed by Fraunhofer was limited only by the resolving

power (Chap. LVIII) of his prism. The actual number of such lines seems to be unlimited, since every improvement in the resolving power of prisms or gratings reveals a greater number of them.

466. Total, Mean, Partial and Relative Dispersion. If sunlight be passed through three prisms having the same refracting angle, one of flint glass, one of crown glass, and one a hollow prism with plane glass sides and filled with water, the resulting spectra will be found to differ greatly in length. Thus (Fig. 265), the spectrum from the flint glass prism is about twice as

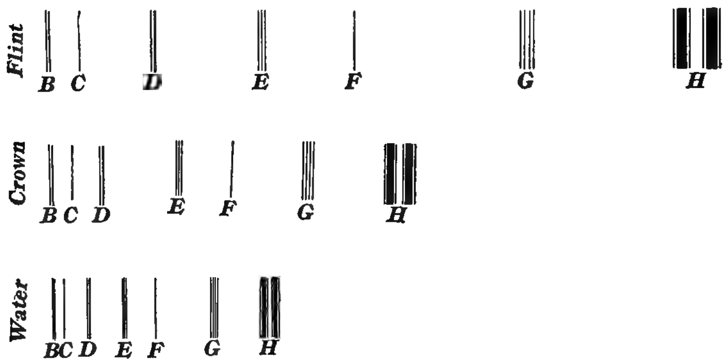


FIG. 265.

long as that from the one of crown glass, and three times as long as that from the water prism. It is clear that the various colors undergo widely different deviations by prisms of the same angle but of different substance. A careful study of the dispersion of various refracting media is therefore a prerequisite for the scientific construction of optical instruments.

If μ_B, μ_C, ... μ_H represent the refractive indices of a given prism for the corresponding Fraunhofer lines, then $\mu_H - \mu_B$ represents the *total dispersion* for that prism in the spectral region from B to H. Likewise the differences $\mu_C - \mu_B$, $\mu_D - \mu_C$, $\mu_F - \mu_D$, etc., are termed the *partial dispersions* for the corresponding regions. Since the middle region, C to F, includes the brightest part of the spectrum, the difference $\mu_F - \mu_C$ is termed the *mean dispersion*. Also, since the D line lies near

the brightest part of the spectrum, its index, μ_D, is called the *mean index* for a given substance. The ratio of the mean dispersion to the quantity $\mu_D - 1$ is termed the *relative dispersion* or *dispersive power*, $\frac{\mu_F - \mu_C}{\mu_D - 1}$, of a substance, and is used to characterize different grades of optical glass with reference to the possibility of selecting achromatic combinations.

From Article 451 we have seen that for prisms of small refracting angle A, the deviation D for any color is defined by the equation

$$D = A(\mu - 1)$$

Obviously the angular dispersion ψ, for any two lines such as B and H, is given by the angular difference between their respective deviations, or,

$$\psi = A(\mu_H - \mu_B) \tag{529}$$

It remains to be noted that since different glasses differ so widely in *relative dispersion*, it is within the power of the optician to produce at will prism combinations that will give *either deviation without dispersion or dispersion without deviation*, according as the need for each may arise. The discussion of these combinations will follow later.

467. Irrationality of Dispersion. If, through variation of the refracting angles, the spectra from the flint glass, crown glass and water prisms should be made of exactly the same length, it would be found that the colors are not equally dispersed by the three substances. If the three spectra be arranged one above the other, and so adjusted that the B and H lines coincide, it is immediately evident (Fig. 266) that the spectra are quite different.

Thus in the water spectrum, the interval from B to F is equal to that from F to H. In the crown glass spectrum the interval from B to F is smaller than that from F to H, and in the flint glass spectrum it is smaller still. It thus appears that in flint glass the dispersion is relatively greater in the blue end of the spectrum, while for water the dispersion is relatively greater in the red end.

If prisms of other substances should be examined, still other variations in the dispersion of the various colors would be found. In short *there is no connection* between the change in refractive index with the wave length *in different substances.* Should the law of this change be known for some one substance, such a law would give us no information as to the change to be expected in another substance. This peculiarity is termed the *irrationality of dispersion.* It is due to this fact that the spectra produced by different prisms cannot readily be compared with each other. Herein lies a fundamental difference between *prismatic spectra*

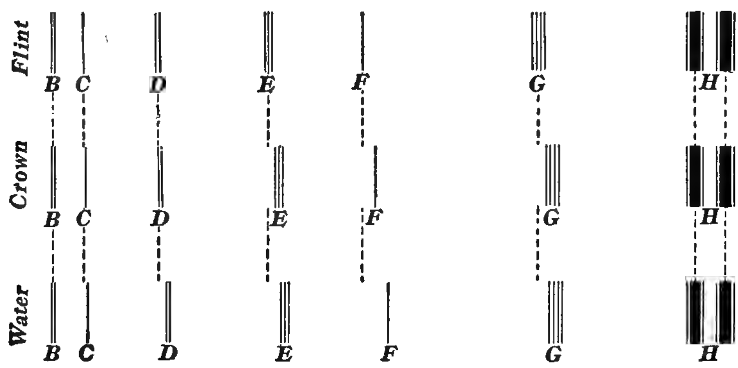


FIG. 266.

and spectra produced by gratings. In grating spectra, as we shall see later, the deviation for every color and for every grating is directly proportional to the wave length. Hence any grating spectrum may at once be compared with any other. On this account the spectra from gratings are termed *normal spectra.*

468. Anomalous Dispersion. A more surprising peculiarity of dispersion was discovered by Christiansen in 1870, who found for a hollow prism filled with an 18.8 per cent solution of fuchsine, or aniline red, an entirely new succession of colors in the spectrum. In this case the violet light was refracted the least, then came the red and then the yellow; the green and bluish green were wholly absorbed. Such a spectrum is termed *abnormal,* and the substance is said to exhibit *anomalous dispersion.*

Through the investigations of Kundt, it has been demonstrated that anomalous dispersion is characteristic of all substances possessing "surface color," *i. e.* substances which exhibit a different color by reflected light from that shown in transmitted light. Thus fuchsine is reddish violet by transmitted light, but an intense green by reflected light. All bodies of this class exhibit *total reflection for certain colors of the spectrum at all angles of incidence*, and their solutions, even when very dilute, show marked absorption bands in these same colors.

Kundt observed anomalous dispersion in the solutions of a large number of substances, among which were the red, blue, green and violet anilines, carmine, indigo, orsellin, litmus, hematine, chlorophyl, iodine in carbon disulphide, and the extracts of red wood and sandal wood. Among solids, cobalt blue glass shows anomalous dispersion very well.

It was theoretically proved by von Helmholtz, and subsequently confirmed by the experimental work of Pflüger, Wood and Magnusson, that in all cases where a substance presents one or more absorption bands in its spectrum, the refractive index of the substance for colors *immediately below the absorption band is increased enormously*, while the index for colors just above the band is correspondingly *decreased*. The index for the color that is wholly absorbed becomes infinite. From this point of view the apparent "anomaly" in the dispersion of strongly absorbing substances disappears. Fuchsine absorbs the green, hence the refractive indices of the red and yellow are increased and that of the violet diminished; the violet is therefore least refracted and the yellow the most. It is now well established that anomalous dispersion is also exhibited by all gases and vapors showing absorption bands in the spectrum.

469. Chromatic Aberration. Since it has been shown that the refractive index of a substance differs for different colors, being greatest for violet and least for red, it is clear from the formula

$$\frac{1}{f} = (\mu - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$$

that the focal length of a lens is shortest for the violet and longest for the red light. In other words, if a simple lens (Fig. 267) be presented to a parallel beam of white light, the violet rays will come to a focus first, at v, and the red last, at r, the foci for the remaining colors being distributed along the axis between these points. In consequence of its dispersion, therefore, a single lens, instead of producing for each point object a *single point image*, in reality produces a *series of colored images*. Hence the image of any object formed by a single lens in white light is surrounded by fringes of color and is generally blurred and indistinct. This defect of lenses is called *chromatic aberration*. An image in which all the colors are focused at the same point is *colorless* or *achromatic*.

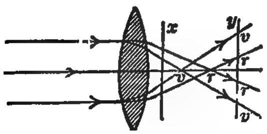


FIG. 267.

About 1670 Newton concluded, from his experiments upon the refractive indices of glass and water, that the relative dispersion of all substances *were the same*, and he consequently despaired of being able to make a lens which would give a colorless image. It was not until 1758 that the London optician, John Dollond, succeeded in making such a lens.

We have seen that light which has been dispersed by one prism can be reunited into an achromatic image by a *second, similar* prism with its refracting edge turned in the opposite direction. In this case the second prism corrected *both the deviation and the dispersion*. If now, the second prism be made from a glass having a *high relative dispersion*, as flint glass, then the dispersion of the crown glass prism may be corrected by a flint glass prism of *much smaller angle*, and an *outstanding deviation* will thus be left. The condition for achromatism in the two prisms is obtained by equating their angular dispersions (529), or

$$A(\mu_H - \mu_B) = A'(\mu'_H - \mu'_B) \tag{530}$$

whence

$$\frac{A}{A'} = \frac{\mu'_H - \mu'_B}{\mu_H - \mu_B} \tag{531}$$

or, *for achromatism the angles of the two prisms must be inversely proportional to the dispersions of the two prisms in the region to be achromatized.*

In an exactly similar way it is possible so to combine a convex lens of crown glass and a concave lens of flint glass that their resulting dispersion shall be zero and the system shall still have any desired focal length. In *lenses* the *curvatures* of the four surfaces are varied, as the *angles* have been in the case of the two *prisms*, but since there are *four* radii to be chosen, two may be adjusted for achromatism and the remaining two so selected as to reduce spherical aberration to a minimum.

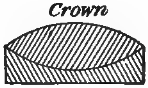


FIG. 268.

Such a combination (Fig. 268) is called an *achromatic lens*, and is said to be corrected for some two colors. If the lens is to be used for visual work, it is usually corrected for the region C to F. If the combination is to be applied to photography, the achromatization is computed for the region G to H, since in this region the chemical action of light is most pronounced.

Owing to the irrationality of dispersion, it is obviously impossible to achromatize any region perfectly unless the "law of change" of the dispersion throughout the spectrum of the two glasses is the same over that region. The slight residual color defects in an achromatic pair are termed *secondary spectra*, and may be reduced still further by the addition to the combination of a third lens of the proper curvature and dispersion. By the use of the new optical glasses introduced by Abbé and Schott of Jena, it is now possible to achromatize more perfectly with two lenses than it had formerly been possible to do with three.

470. Direct Vision Spectroscope. If two prisms, one of flint and the other of crown glass, have their angles so chosen that their *deviations* for some one color shall be equal, that is, so that

$$A(\mu - 1) = A'(\mu' - 1) \tag{532}$$

there will remain an *outstanding dispersion* due to the flint prism. This fact has been utilized in the construction of the

direct vision spectroscope. Usually the instrument contains two ninety-degree, flint glass prisms combined with three prisms of crown glass.

Fig. 269 shows a combination of one dense flint prism P, and two crown glass prisms. By means of the achromatic lens O, the slit S is sharply focused for sodium light. On turning the instrument toward some source of white light, the image of the slit is spread out into a spectrum. The width of the slit S is adjustable, and by making it narrow the Fraunhofer lines are readily seen in ordinary daylight. Light from the source under examination is admitted directly into the instrument by the slit (ray 2), while by means of the totally reflecting prism V (Art. 451), and the mirror D, light from a second source (ray 1) may be introduced, and the spectra of the two sources observed side by side. The deviation of the prism train is usually adjusted to zero for the D line.

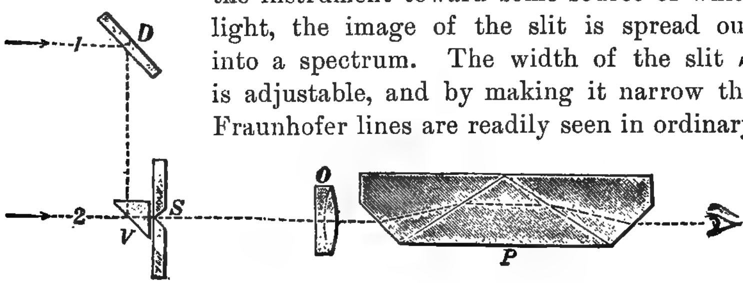


FIG. 269.

Problems

1. If the indices of refraction for the four substances, flint glass, crown glass, water and carbon disulphide are

	B	C	D	F	H
Flint glass	1.6127	1.6144	1.6193	1.6315	1.6527
Crown glass	1.5301	1.5311	1.5339	1.5404	1.5509
Water (18°.7 C)	1.3310	1.3320	1.3336	1.3380	1.3448
CS_2 (18°.7 C)	1.6182	1.6219	1.6308	1.6555	1.7020

compute the angle of a flint glass prism necessary to achromatize the region from B to H in a crown glass prism whose angle is 3° 30′. *Ans.* 1° 49′ 12″.

2. From above table compute the linear separation of the lines C and F produced by a hollow prism filled with carbon disulphide, of angle 40°, if the slit image be half an inch wide, and the prism be 30 ft. from the screen. *Ans.* About 8.2 in

3. In a direct vision spectroscope (Fig 269), the flint glass prism has an angle of 54°. Compute the base angles of the two crown glass prisms needed to give direct vision for the D line, using data given above. *Ans.* 31° 21′.

4. Compare the lengths between B and H, of spectra produced by two hollow prisms of angle 40°, one filled with water and the other with carbon disulphide. Prism distant 20 ft. from screen. *Ans.* (*a*) 2.313 in.
(*b*) 14.02 in.

5. AB is the diameter of a polished semicircular arc APB. A ray of light proceeds from a point Q in the tangent at A, and after reflection at P and B returns to Q. Show that if the length of the ray's path be 2 ft, the mirror's diameter is very nearly 7.35 in.

6. A person whose height is h observes vertically beneath his eye an object at the bottom of a clear pool: he then removes to a distance d, keeping his eye on the object, when his line of vision makes an angle of 45° with the surface; show that if $\mu^2 = 2.5$, the depth of the pool $= 2\ (d - h)$.

CHAPTER LIV

OPTICAL INSTRUMENTS

471. Projection Apparatus. The projection lantern (Fig. 270) consists essentially of two optical systems, one for condensing light upon the object, and the other for forming the image upon the screen. The source of light is usually either the limelight or the arc lamp. This source A is placed just outside the focus of the condenser C, which generally consists of

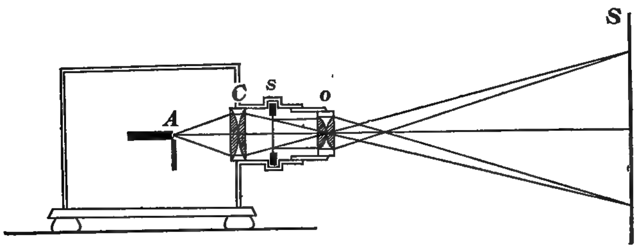


FIG. 270.

two plano-convex lenses of short focus with their convex surfaces turned toward each other. The cone of light emerging from the condenser falls upon the object, a transparent drawing or lantern slide, and then passes through the projecting lens or objective o. This is usually an achromatic lens of about one foot focal length. Since the object is just outside the focus of the objective, the image projected upon the screen is real, inverted and magnified.

472. The Camera Obscura. In the *camera obscura* we have the projection apparatus reversed. It consists, as its name indicates, of a darkened chamber having an opening on one side, in front of which is placed a projecting lens. Upon the wall

of the chamber opposite the lens is a white screen for receiving the image of illuminated objects which chance to be in front of the lens. The image is real, inverted and smaller than the object. This instrument, in a portable form, was formerly used as an aid in sketching buildings or the prominent features of a landscape. Since the discovery of the various methods of fixing an image upon a photographic plate by the chemical effects of light, it has developed into the photographer's camera.

The simple lens has been replaced by a more or less elaborate optical system carefully corrected for spherical and chromatic aberrations. The sides of the chamber have been made flexible to admit of focusing the image accurately, and the size of the aperture is carefully adjusted to the existing degree of illumination by a series of graded diaphragms. In its present form, the photographer's camera is one of the most delicate, accurate and important of optical instruments, revealing many things that escape the eye, since it secures the cumulative effect of faint luminous impressions.

473. The Eye. From an optical point of view, the eye is a photographic camera with an automatic adjustment for focus, and a sensitive plate that reports the images directly to the brain. Fig. 271 represents a sectional view of the human eye. The *cornea* *a* and the *sclerotic* coat *b* form the outer coating of the darkened chamber, a nearly spherical cavity about one inch in diameter. The sclerotic coat is tough and strong, holding

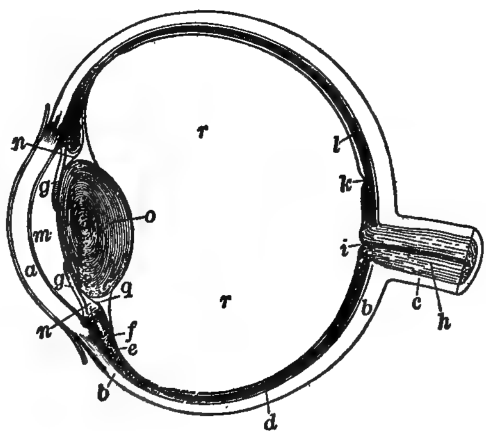


FIG. 271.

the ball of the eye in shape, and forms the "white of the eye." The cornea is a transparent, hornlike membrane, which serves

to admit light into the eye and to retain the *aqueous humor m* in a lenticular form.

The second coat *d* is the *choroid* coat, a black, opaque membrane which serves to darken the interior of the eye. In front, the choroid coat gives place to the *iris gg*, an opaque, colored diaphragm, which regulates the amount of light entering the eye through the circular opening in its center, which is called the *pupil* of the eye. Just back of the iris lies the *crystalline lens o*, a double convex lens of unequal curvatures, composed of a transparent, jelly-like substance built up in concentric layers. The optical density of these layers increases as we go toward the center. The interior cavity *rr* of the eye is filled with a viscous mass of fluid of the highest transparency, called the *vitreous humor*.

The third coat is the *retina l*, a delicate membrane formed by the expansion of the optic nerve, which serves as the sensitive screen for the reception of the images formed by the optical system of the eye. The inner surface of the retina on the posterior part of the eye is characterized by a mosaic-like structure of nerve endings, termed *rods* and *cones*. These microscopic little bodies stand perpendicular to the retinal surface, are closely crowded together, and each one is furnished with a separate nerve fiber. These are believed to form the real *light sensitive layer* of the retina.

The retina possesses two spots of peculiar interest. The "*yellow spot*," situated in the optical axis of the eye, marks a small area where the retinal layer is slightly thickened. In its center is located the *fovea centralis k*, where the sensitive layer is reduced to an exceeding thinness, only the most minute nerve filaments extending into it. This is the spot of *finest space discrimination*. It is upon this spot that the image of anything is brought which we wish to examine attentively. The point *i* at which the optic nerve enters the eye is entirely insensible to light, and is called the "*blind spot*."

The adjustment of the crystalline lens for focusing objects at varying distances is accomplished by changing the thickness of the lens itself, through the action of the *ciliary muscle e*.

Recent investigations show that this muscle by contracting thickens the lens, and thereby renders more accurate the focus for objects near at hand. This power of adaptation, or of *accommodation* as it is called, is most marked in young children, and gradually diminishes through life.

A normal eye can accommodate for distinct vision for objects at all distances from 6 inches from the eye (near point), up to infinity (far point). Within this range of accommodation, however, there is a point whose distance from the eye is termed the *distance of distinct vision.* It is defined as the distance at which a normal eye can most readily read ordinary print, and is assumed to be *25 centimeters* or *10 inches.*

Since the formation of images by the eye is identical with that of the photographic camera, it follows that the images perceived by the retina *are inverted.* The ability of the eye and brain to interpret inverted images as belonging to erect objects is probably gained by experience, through combination with other sense perceptions, especially with that of touch.

474. Defects of Vision. Through structural defects the eye may be unable to accommodate for distant objects. The lens in this case is too powerful and the image is focused at *f*, *in front* of the retina, instead of *upon it.* A child so afflicted instinctively brings any object nearer and nearer to its eye. thereby causing the image of the object to recede farther and farther from the crystalline lens and finally to fall upon the retina, when distinct vision results. In cases of this kind the distance of distinct vision is frequently not more than three or four inches. Such an eye is called a *short-sighted* or *myopic eye.*

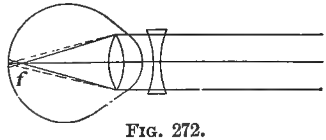


FIG. 272.

The myopic eye requires a concave lens to counteract the excessive refraction of the eye, in order to bring the images of distant objects within the range of accommodation of the eye, as shown (Fig. 272).

Again, the power of the lens may be below the normal, in which case the focus for parallel rays lies *behind* the retina

instead of upon it. Such an eye is termed *hypermetropic* or *far-sighted.* For such an eye a convex lens is needed to enable the eye to bring the images of distant objects upon the retina (Fig. 273).

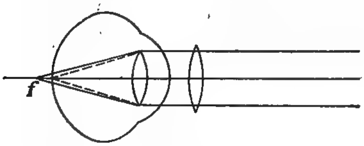


FIG. 273.

The eye gradually loses its power of accommodation during life, and at about the forty-fifth year it becomes unable to accommodate for objects near at hand, although the vision for distant objects is as good as ever. Such an eye is said to be *presbyopic.* For such eyes convex glasses are to be used for reading and for examining objects close at hand.

It frequently happens that the eye has different power in different planes. A bright point is seen not as a *point* but *as a line.* Horizontal, vertical or diagonal lines, like the spokes of a wheel or the letters on a clock dial, may be seen with unequal distinctness. Such a defect is called *astigmatism.* For such eyes glasses having cylindrical surfaces are used.

475. Apparent Size and Magnification. The angle subtended by an object at the optical center of the eye is called the *visual angle.* The apparent size of a linear object is measured by the visual angle which it subtends. Thus if d be the distance of an object from the eye, b the distance of the retinal image from the optical center of the crystalline lens, δ and β the lengths of object and image respectively, then V, the visual angle, or the apparent size of the object, is

$$V=\frac{\delta}{d}=\frac{\beta}{b} \qquad (533)$$

Hence *the apparent size of a linear object is inversely proportional to its distance from the eye.* Also the size of the retinal image of an object is inversely proportional to the distance of the object from the eye, and directly proportional to the size of the object.

The clearness with which the minute details of an object can be distinguished increases with the size of the retinal image

Hence to see a thing clearly we bring the object up to the near point of the eye and thus secure the maximum efficiency of the unaided eye. The effect of any optical instrument which may be used to assist the eye consists simply in increasing the apparent magnitude of the object. Hence the *magnifying power of any instrument* may be defined as the ratio of the apparent magnitude *through* the instrument, to the apparent magnitude *without* the instrument, or

$$\frac{\textit{Apparent magnitude through instrument}}{\textit{Apparent magnitude without instrument}} = \textit{Magnification} \qquad (534)$$

476. The Simple Microscope. The simple microscope is a converging lens of short focus. When an object is placed slightly nearer to the lens than its focal distance, an eye brought close up to the lens perceives a virtual, erect and magnified image $A'B'$, as shown in Fig. 274. If we conceive the eye to be accommodated for infinite distance, then the object is moved up to the focus F, and the image moves off to infinity. The object is now seen clearly by the eye although much nearer to the eye than would have been possible for distinct vision except for the lens. The microscope thus enables the eye to see an object clearly at a much shorter distance, and hence acts as an aid to the accommodation for this distance.

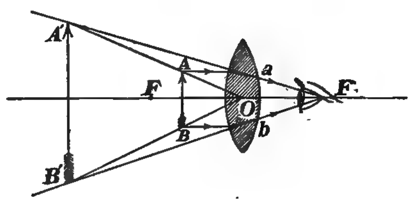


FIG. 274.

The visual angle under which the object appears *through the lens* is $\frac{AB}{f}$, where AB is the length of the object and f is the focal length of the lens. The visual angle subtended by the object *without the lens* is $\frac{AB}{25 \text{ cm}}$; hence the magnification m, produced by the lens for an eye adjusted for infinity, is

$$m = \frac{AB}{f} \div \frac{AB}{25 \text{ cm}} = \frac{25 \text{ cm}}{f} \qquad (535)$$

or *the magnification due to a simple lens is obtained by dividing the distance of distinct vision by the focal length of the lens.*

To a certain degree the effect of magnification may be obtained without the use of a lens, by looking at a brilliantly illuminated object through a very small hole, as through a pinhole in a card. The hole acts as a diaphragm or second pupil to the eye, and thus, by stopping off the rays from the edges of the pupil, permits of the formation of an image at a distance much less than the normal distance of distinct vision, 25 cm. Such magnification is obtained, however, at the cost of illumination, while the lens accomplishes the same result without reducing the illumination.

477. The Astronomical Telescope. In the astronomical telescope (Fig. 275), the objective L is a converging lens of long focal length F, which forms a real, inverted image A_1B_1 of a

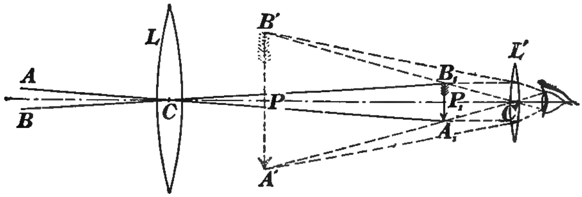


FIG. 275.

very distant object. This image subtends the same angle at C, the center of the objective, as does the object itself. Hence $\frac{A_1B_1}{F}$ represents the apparent magnitude of the object.

The real image formed by the objective is viewed through the eyepiece L', used as a simple magnifier. The eyepiece forms a virtual, magnified image $A'B'$, of the image A_1B_1. The apparent magnitude of the object as seen *through the telescope* is, therefore, the visual angle $B'C'A'$, or $\frac{A_1B_1}{f}$, where f is the focal length of the eyepiece. The magnification m of the telescope is therefore given by the equation

$$m = \frac{A_1B_1}{f} \div \frac{A_1B_1}{F} = \frac{F}{f} \tag{536}$$

or *the magnification of the telescope is obtained by dividing the focal length of the objective by the focal length of the eyepiece.*

478. The Compound Microscope. The compound microscope in its simplest form may be regarded as a telescope used for examining objects near at hand. The object glass O (Fig. 276) is a converging lens of very short focus placed near the object AB, but still outside its principal focus. The real, inverted image $A'B'$ is much magnified. This image is viewed by the eyepiece E, also a converging lens, used as a simple magnifier, which forms the virtual, highly enlarged image ab.

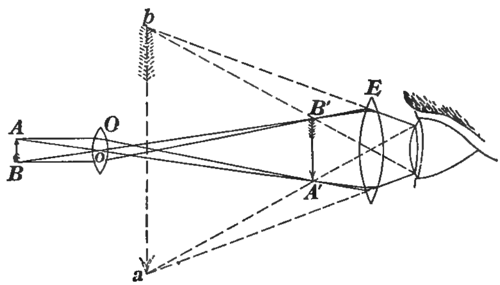


FIG. 276.

The approximate magnification may be written down at once. Letting o represent the center of the objective, we have for the magnification of the object glass,

$$\frac{A'B'}{AB}=\frac{B'o}{Bo}=\frac{L}{F} \tag{537}$$

where L is the length of the microscope tube and F the focal length of the objective. Also, since the magnification by the eyepiece is $\frac{25 \text{ cm}}{f}$, the total magnification is

$$\frac{L}{F}\times\frac{25}{f}=\frac{25\,L}{Ff} \tag{538}$$

The practical advantage of the compound microscope consists in the possibility of using higher powers. With the simple magnifier the working distance soon becomes too small to enable one to observe with comfort or accuracy. Thus, with a magnification of 100, the lens must be brought within 0.1 of an inch of the object, and the eye must approach the lens very closely

indeed if none of the rays are to fall outside the pupil. At the same time the dimensions of the lens become too minute for accuracy in grinding.

479. Spectroscope and Spectrometer. It has been noted that the conditions for obtaining a pure spectrum were approximated by Fraunhofer in his arrangement for observing the dark lines in the solar spectrum. The one disadvantage consisted in the loss of light due to the distance between the slit and the prism. In the modern spectroscope (Fig. 277) this loss is avoided by the use of the *collimator*. The collimator is simply a telescope in which the eyepiece is replaced by an adjustable slit. When put in adjustment, the slit is brought into the focal plane of the achromatic objective, so that all light entering the slit leaves the objective as parallel rays. In this way the full intensity of the light from the slit is preserved and the additional advantage of parallelism of the rays is secured.

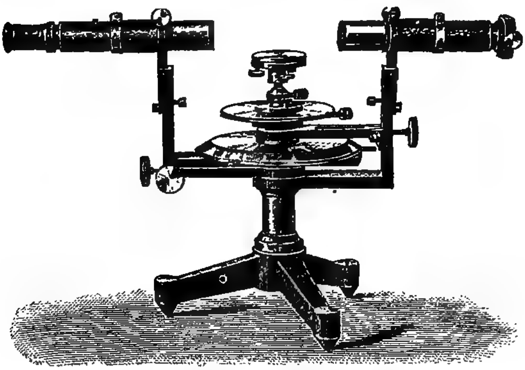


FIG. 277.

The remaining parts of the spectroscope are a suitable table for mounting the prism and an observing telescope with achromatic objective for receiving the spectrum. The telescope moves about an axis concentric with that of the prism table. When the instrument is furnished with a graduated circle for determining accurately the positions of telescope and prism, it becomes a *spectrometer*, and may be applied to a variety of optical measurements.

In the Abbé spectrometer (Fig. 278), the functions of the collimator and the observing telescope are combined. A small,

totally reflecting prism furnished with an adjustable slit is inserted in the focal plane of the eyepiece of the observing telescope, at one side of the field of view. Light admitted through a small window in the side of the telescope tube is reflected by the prism through the slit, directly into the optical

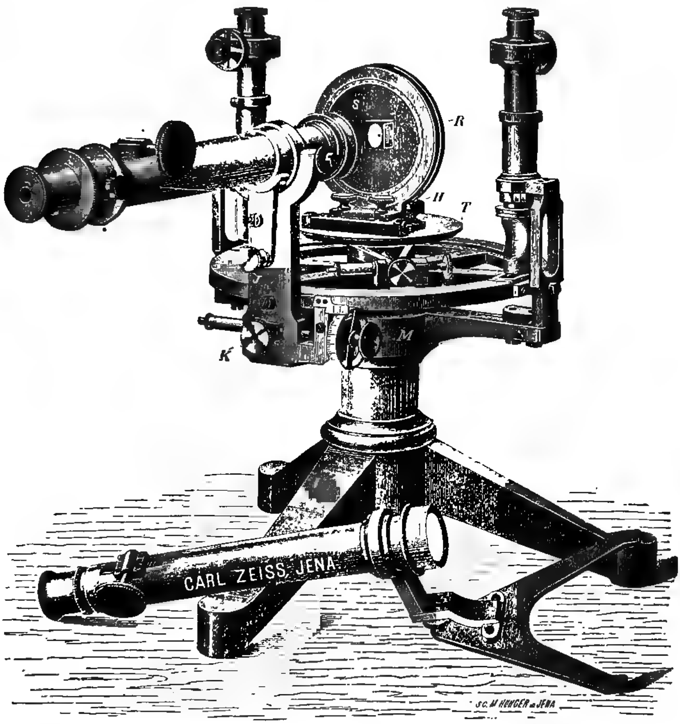


Fig. 278.

axis of the telescope, and leaves the objective as parallel rays. By the use of the Abbé prism described in Article 452, the dispersed light is returned upon its path and the spectrum is viewed by the telescope in the unobstructed part of the field. For the position of minimum deviation the refracted image of the slit is made to coincide with the slit itself, and the settings can be made with great exactness. By means of the micrometer

screw *M*, the dispersion of any substance in the various spectral regions is measured independently of the graduated circle, thus greatly reducing the labor of making the measurements and at the same time securing a marked increase in the accuracy of the results. An auxiliary telescope permits of the use of the instrument in the ordinary form.

PHYSICAL OPTICS

CHAPTER LV

VELOCITY OF LIGHT

480. Velocity of Light — Roemer's Method. The first definite proof of the finite velocity of light was given by Roemer, a Danish astronomer, in 1675, from a study of the eclipses of the satellites of the planet Jupiter. This planet has several satellites which revolve about it as our moon does about the earth. The inner one of these satellites has a mean period of 42 hr 28 min 36 sec. As it passes behind the planet it disappears from view quite suddenly, and the interval between two eclipses can be determined with considerable accuracy. Roemer noticed that this interval was not constant throughout the year; that it was *less* than the period of revolution of the satellite when the earth was *approaching* Jupiter, and *greater* than this period when the earth was *receding* from Jupiter.

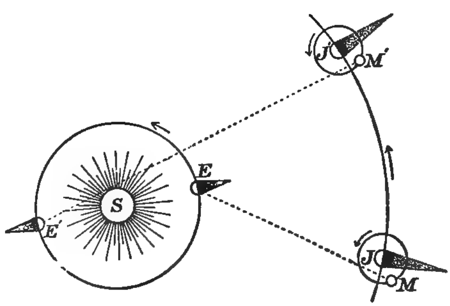


Fig. 279.

Thus (Fig. 279), the period of the satellite coincided with the eclipse interval when the earth and Jupiter occupied the positions $E'J'$, or EJ, but as the earth passed from E' to E, the eclipses occurred systematically *ahead* of the computed time by differences whose sum at E amounted to about 1000 seconds. From E to E' they *fell behind* by the same amount. Roemer concluded that this

difference in time must be required by the light to traverse the difference in path EE' between these two positions. Assuming the diameter of the earth's orbit to be 186,000,000 miles, the velocity of light was computed to be 186,000 miles per second.

Roemer's wonderful discovery was received with little favor by the scientific world and was practically disregarded for over fifty years, until Bradley's discovery of the aberration of light, in 1728, gave an additional method for measuring this important physical constant.

*** 481. Velocity of Light — Foucault's Method.** In 1850 Foucault, a French physicist, described a method for the direct determination of the velocity of light, which in its improved

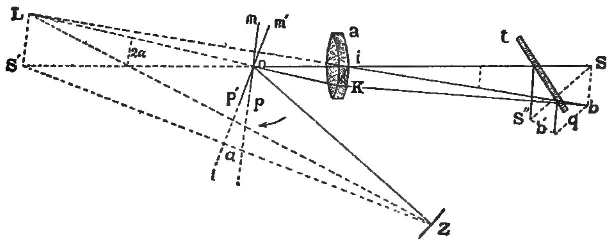


FIG. 280.

form has given the most accurate values yet attained. Foucault's original arrangement consisted of an illuminated slit S (Fig. 280), from which light passed through a lens K, of long focus, to a mirror mp, which was capable of rapid rotation about an axis through o, normal to the plane of the paper. From this mirror mp the light was reflected to a concave mirror Z, striking it normally, since the center of curvature of Z lay in o. The mirror Z and the slit S lay in the conjugate foci of the lens K. The light was therefore returned upon its path and after a second reflection at the mirror mp, fell again upon the slit S. By interposing a piece of plane parallel glass t, at an angle of 45° to the path of the light, a displaced image S'' of the slit was produced for the purpose of observation. When the rotating mirror mp was put in motion, a flash of light fell upon the concave mirror at each revolution, and for rotations of more than

30 per second these flashes blended into a steady image of the slit.

Now, if in the time t, needed for the light to travel the distance $2D$, from the rotating mirror to the concave mirror and back again, the rotating mirror should have turned through a small angle α, then the reflected image of the slit would have been displaced through twice that angle, or 2α. If, therefore, we call the distance from the rotating mirror to the slit r, and the displacement of the slit s, then from a knowledge of the quantities D, s, r, and n, the number of rotations of the mirror per second, the velocity of light V can be calculated. Thus

$$t = \frac{2D}{V} \tag{539}$$

Also, if the mirror turn through a small angle α in time t, then

$$\alpha = \omega t = 2\pi n t \text{ or } t = \frac{\alpha}{2\pi n} \tag{540}$$

Now the light is turned through an angle $2\alpha = s/r$, where the displacement of the slit image is small, and

$$\alpha = \frac{s}{2r} \tag{541}$$

From (540) and (541) we have

$$t = \frac{s}{4\pi n r} \tag{542}$$

and combining (539) and (542) we get

$$\frac{s}{4\pi n r} = \frac{2D}{V} \tag{543}$$

or finally,

$$V = \frac{8\pi n r D}{s} \tag{544}$$

In Foucault's experiment D was 20.24 m, r was 1.0257 m, the speed of the mirror was 400 revolutions per second, and the measured displacement s of the slit image was 0.7 mm. From these values he computed the velocity of light to be

$$V = 298{,}000 \frac{\text{km}}{\text{sec}}$$

In Foucault's experiment the displacement, 0.7 mm, was a quantity too small to be measured with accuracy, while the speed of rotation of the mirror was enormous. In 1878 Michelson so modified the method of Foucault as to render it capable of far greater accuracy. The lens L (Fig. 281) was so placed that all rays from the slit s, after reflection at the rotating mirror m, left the lens as approximately parallel rays. For the concave mirror, a plane mirror m' was substituted, and the distance D was increased to 605 m, instead of 20 m as in Foucault's experiment. In this way Michelson succeeded in securing a displacement of the slit s of 133 mm, where r was 9 m, and the speed of rotation for the mirror was 257 revolutions per second. The mean of Michelson's measurements gave the velocity of light.

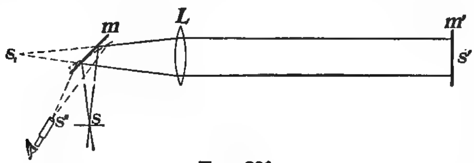


FIG. 281.

$$V = 2.999 \times 10^{10} \frac{\text{cm}}{\text{sec}}$$

Newcomb, also using the method of the rotating mirror, obtained a value in close agreement with this.

The experiment of Foucault was originally designed for the purpose of actually measuring the velocity of light in media of different density in order to decide experimentally between the corpuscular and the undulatory theories of light. According to the corpuscular theory, light should travel faster in a dense medium than in a rare one; from the principles of the undulatory theory exactly the opposite conclusion was reached. Experiment alone could decide. Foucault in 1850 measured the velocity of light in air and in water and found the velocity in water to be less than in air. Since that time the emission theory of light has been definitely abandoned, and the undulatory theory in some form or other is now generally accepted among scientific men. A statement of the theory forms the topic of the next article.

* **482. Undulatory Theory of Light.** According to the undula tory theory, light consists in an extremely rapid periodic change of condition which is transmitted from point to point in the form of transverse waves. Since experience shows that light traverses a space the more readily the less ponderable matter the space contains, it is concluded that light may be propagated even in space containing no ponderable matter, *i.e.* in a vacuum. It is assumed that the universal medium for the transmission of luminous disturbances is the ether, which consequently cannot be ponderable matter, although it must possess many properties in common with it.

In order to formulate the undulatory theory it is necessary to assume that the ether fills all space, and that it has different properties in different media. The two most important variations of this theory are the *mechanical* theory and the *electromagnetic* theory of light.

According to the mechanical conception, light is assumed to be due to a vibratory motion of ether particles arising from definite displacements of these particles from their positions of equilibrium. In the elaboration of this theory the laws of elasticity as manifested in ponderable matter are assumed to hold in the ether without modification. In following out the consequences of these assumptions it is necessary to bear in mind certain results previously established for wave motions in general.

(*a*) When, in an elastic medium of density d and coefficient of elasticity e, a molecule is displaced, the general equilibrium of the medium is destroyed; all the neighboring molecules experience a movement which is propagated from point to point in all directions with a velocity

$$V=\sqrt{\frac{e}{d}} \tag{545}$$

This same law is held to apply to the ether where V is the velocity of light.

(*b*) The intensity of any vibratory disturbance proceeding from a point source is *directly proportional* to the *square of the amplitude*.

According to the electromagnetic view, light is assumed to be due to the propagation in space of periodic changes of the electrical and magnetic intensities in the dielectric, such as accompany the oscillatory discharge of a condenser. The fundamental assumption of this theory is that the velocity of light in a non-absorbing medium is identical with the velocity of an electromagnetic wave in the same medium. The displacement currents of Maxwell (Art. 315) are assumed to be accompanied by magnetic displacements at right angles to the electric field, similar to those manifested by ordinary currents. The exact character, however, of these displacements remains undefined, and for this reason the mechanical concept of the undulatory theory will be adopted in this text, since it presents fewer difficulties in an elementary presentation.

***483. Equations of Wave Motion.** If, in place of an infinitely small and instantaneous movement, a particle of ether execute regular vibrations, its oscillations, if simple harmonic, may be expressed by the equation

$$y = A \sin \frac{2\pi t}{T} \tag{546}$$

in which y is the displacement of the ether particle at any time t, A the amplitude and T the period.

If light be transmitted with a velocity V, from an ether particle P_1 to a second particle P_2 distant x from P_1, the time required for transit is x/V. Now if equation (546) represent the condition at P_1, then the condition at P_2 is represented by

$$y' = A' \sin\left(2\pi \frac{t - x/V}{T}\right) \tag{547}$$

since P_2 is always in a given condition of vibration, x/V *seconds after* P_1 has been *in the same condition*. The difference in the two vibrations, as may be readily seen, is a difference in *amplitude* and a difference in *phase*. If the wave be generated at a point source in a homogeneous and isotropic medium, then the disturbance travels outward in all directions with the same

velocity. Hence all points on the surface of a sphere with center at P, and radius x, must be in the *same phase*. Surfaces containing only points in the *same phase of vibration* are called *wave surfaces*. The distance from one wave surface to the next surface having the *same phase of vibration* is called a *wave length* λ, and is defined in terms of velocity and period of vibration by

$$\lambda = VT$$

If we introduce λ into equation (547), we have for the condition at P_2,

$$y' = A' \sin 2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right) \tag{548}$$

***484. Superposition of Small Vibrations.** Whenever an ether particle is actuated at the same time by impulses due to two sets of vibrations, the resultant motion is that due to the *superposition* of the *two vibrations*. Since in all cases the amplitudes of vibrations are assumed to be small, the treatment of the problem is termed the superposition of *small* vibrations. If we represent the two impulses actuating the ether particle by the two equations

$$\left.\begin{aligned} y_1 &= A_1 \sin \frac{2\pi t}{T} \\ y_2 &= A_2 \sin\left(\frac{2\pi t}{T} - \frac{2\pi x}{\lambda}\right) \end{aligned}\right\} \tag{549}$$

and for convenience set

$$\frac{2\pi t}{T} = \alpha, \text{ and } \frac{2\pi x}{\lambda} = \beta,$$

then the resultant motion of the ether molecule may be represented by

$$Y = y_1 + y_2 = A_1 \sin \alpha + A_2 \sin(\alpha - \beta) \tag{550}$$

Expanding the right-hand member and adding, we have

$$Y = (A_1 + A_2 \cos \beta) \sin \alpha - A_2 \sin \beta \cos \alpha \tag{551}$$

or putting and

$$\left.\begin{aligned} A_1 + A_2 \cos \beta &= A \cos \phi \\ A_2 \sin \beta &= A \sin \phi \end{aligned}\right\} \tag{552}$$

we get

$$Y = A \sin(\alpha - \phi) \tag{553}$$

This shows that the resultant motion of the ether particle is still a simple harmonic motion, with an amplitude A. The resultant amplitude may be found at once from equations (552) by squaring and adding, whence

$$A^2 = A_1^2 + A_2^2 + 2\,A_1A_2 \cos \beta \tag{554}$$

or, replacing β by its value $\frac{2\,\pi x}{\lambda}$, we have

$$A^2 = A_1^2 + A_2^2 + 2\,A_1A_2 \cos \frac{2\,\pi x}{\lambda} \tag{555}$$

This equation shows that the square of the resultant amplitude, and hence the *resultant intensity*, due to the superposition of two vibrations, depends upon the quantity x, that is, upon the *difference in path* traversed by the two component vibrations. If $x = 0$, or *an even multiple of* $\lambda/2$, then the resultant amplitude is the *sum* of the amplitudes; if $x = \lambda/2$, $3\,\lambda/2$, $5\,\lambda/2$, etc., then the resultant amplitude is the *difference* of the component amplitudes. These results are applied in the theory of interference.

485. Law of Reflection of Light deduced from Huygens's Principle. Let AB (Fig. 282) be a plane wave front meeting the reflecting surface CD, in the direction indicated. In accordance with Huygens's principle (Art. 116), all points on the reflecting surface become new centers of disturbance, as the successive points in the wave front reach them. Consequently the point A, being the first disturbed, acts as a center from which the disturbance spreads backward into the original medium with the original velocity, and in the

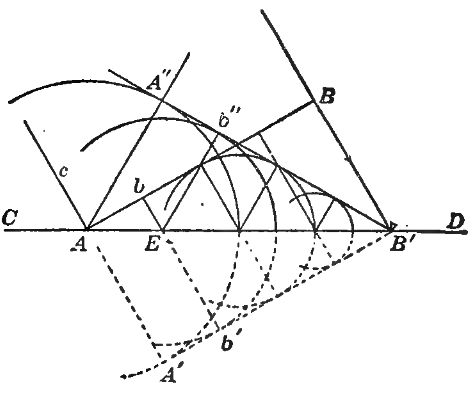


Fig. 282.

time t needed for the disturbance at B to reach B' it has spread through a hemisphere with radius AA' equal to BB', of which the full line above CD is a trace on the plane of the paper.

In the same time the disturbance at b on the incident wave front has reached E, from which point the new wave spreads back in the hemisphere whose radius is Eb'. In a similar way every point on the reflecting surface has generated new waves. The radius for the wave from B' is of course zero at the end of the time t. Now a tangent from B' upon the circle $A''A$ will touch all the other subsidiary circles; consequently the reflecting surface CD has given rise to the new wave front $B'A''$, which now moves off parallel to itself in the original medium.

By the undulatory theory the angle of incidence is the angle included between the *incident wave front* and the *surface*, *i.e.* the angle BAD; similarly the angle of reflection is the angle $AB'A''$ included between the *reflected wave front* and the *reflecting surface*. The equality of these angles can be shown at once from the equality of the triangles ABB', $AB'A'$, and $AB'A''$.

486. Law of Refraction of Light deduced from Huygens's Principle. In the case of refraction of light it may be assumed that the ether imbedded between the material particles of the different media will seem to have different densities and hence will transmit the luminous disturbance with different velocities. Thus suppose the plane wave front AB (Fig. 283) in medium I to meet the plane surface AC of medium II in the direction BC. Now if the second medium should transmit the luminous impulses with the same velocity V as the first, then CD would be the resulting wave front; that is, there would be no refraction at the surface of separation.

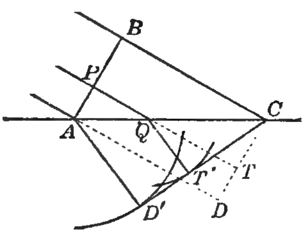


FIG. 283.

But if the second medium transmit light more slowly than the first, say with a velocity V', then in time t while the light in the first medium is passing from B to C, the light in the

second medium will have spread out into a hemisphere about A as a center, and having a radius AD', such that

$$\frac{AD}{AD'}=\frac{V}{V'}, \text{ or } AD'=\frac{V'}{V}\cdot AD \tag{556}$$

The circular arc through D' about A, having its center at A, represents the trace of this hemisphere upon the plane of the paper. In the same way from the point Q, there spreads out a hemispherical disturbance which at the end of the time t has reached a radius

$$QT'=\frac{V'}{V}\cdot QT \tag{557}$$

A tangent from C upon the wave front about A as a center, is also tangent to the entire system of subsidiary waves, and DC is therefore the refracted wave front, which now moves on in the new medium parallel to itself. By a previous definition the angle of incidence is the angle BAC or its equal ACD. Similarly the angle of refraction ACD' is the angle included between the *refracted wave front* and the *refracting surface*.

Applying the law of refraction,

$$\frac{\sin i}{\sin r}=\frac{\sin ACD}{\sin ACD'}=\frac{\frac{AD}{AC}}{\frac{AD'}{AC}}=\frac{AD}{AD'}=\frac{V}{V'}=\mu \tag{558}$$

It thus appears that *when light passes from a medium A to a medium B* the *relative index of refraction for the two media is equal to the ratio of the velocities of light in A and B.*

Problems

1. If the nearest distance for distinct vision for a far-sighted person is 2 ft 11 in, what should be the focal length of the spectacles he would require for reading? *Ans.* $f=14$ in.

2. If the greatest distance of distinct vision for a myopic eye is 3.9 in, what should be the focal length of the proper reading spectacles? *Ans.* $f=6.4$ in.

3. A converging lens placed at a distance of 5.2 cm from a luminous object forms an image on a screen. When the lens is moved a distance of 23 cm nearer the screen, another image is formed. What is the focal length of the lens? *Ans.* $f = 4.4$ cm.

4. A ray of light strikes a plane parallel plate of glass of index of refraction 1.5, at an angle of 70°, and emerges from the other side parallel to its original direction, but displaced laterally through a distance of 5 mm. How thick is the glass? *Ans.* 7.52 mm.

5. A man 6 ft tall stands upon a smooth, level walk 300 ft distant from an observer. A boy 4 ft tall standing between them seems to be of the same height as the man. How far away is the boy if the eye of the observer be in the plane of the walk? *Ans.* 200 ft.

6. A telescope whose objective has a focal length of 12 ft is furnished with two eyepieces whose focal lengths are 1 in and 0.5 in respectively. Compute the magnifying power in each case. *Ans.* (*a*) 144.
(*b*) 288.

7. A person of normal vision uses a convex lens whose focal length is 2.5 cm as a simple magnifier. What is the magnification produced?
Ans. 10.

8. A lens of 0.5 in focal length is used to project a microscopic object upon a screen 30 ft distant from the object. Determine the magnification and the position of the lens. *Ans.* (*a*) 719+.
(*b*) 0.5 in. from object.

9. A double convex lens of radii 80 cm is made from crown glass whose refractive indices are given on p. 506. How far apart upon the axis of the lens are the foci for the *B* and the *F* lines? *Ans.* 1.44 cm.

10. A projecting lantern is to be used in a lecture room, where the screen is 60 ft distant. If the required magnification be 20, find the focal length of the objective needed. *Ans.* 2.875 ft.

CHAPTER LVI

INTERFERENCE

487. General Statement. Thomas Young showed that under certain circumstances two nearly parallel beams of light do not, when superposed, produce increased illumination, but that they may even so disturb each other's effects as mutually to extinguish each other and produce darkness. In such cases the light waves are said to *interfere*, and the resulting phenomena are classed under the head of *interference phenomena.*

Interference phenomena are of two general kinds: first those, in which the two interfering pencils have undergone only reflections and refractions, and have had certain phase differences produced thereby; the second, in which interference takes place between subsidiary waves starting from *different parts of the same wave front.* The latter phenomena are usually classed as *diffraction phenomena.*

It was shown in the study of sound that two sets of sound waves might be made to interfere and produce either a sound of increased intensity or total silence, according as the two vibratory motions were in the *same* or *opposite* phase. Similar phenomena may be produced in the case of two trains of light waves *provided the two sets of waves proceed originally from the same source.* If the two pencils proceed from two different sources, as from two flames or from different parts of *one and the same flame*, they are incapable of producing interference. Such sources are termed *incoherent.* Two pencils proceeding from the same point source are termed *coherent pencils.*

If two sources are to produce interference, their phases must always be either exactly the same, or they must have a fixed difference in phase. In the case of incoherent pencils, the

sources may change their difference in phase many thousands or even millions of times in a second, and yet the wave trains emitted during each interval may include millions of individual waves. A simple calculation may serve to make this clear. Thus from the established value for the velocity of light $V = 3 \times 10^{10} \frac{\text{cm}}{\text{sec}}$ we may deduce the frequency of vibration for light of any wave length. For sodium light, whose mean wave length is 5893×10^{-8} cm, we have

$$\lambda = VT = \frac{V}{n}$$

whence $$n = \frac{V}{\lambda} = \frac{3 \times 10^{10}}{5893 \times 10^{-8}} = 5 \times 10^{14} \textit{ vibrations per second.}$$

If we conceive light to be due to periodic disturbances in the ether occasioned by collisions of the molecules in a heated substance, then these changes in phase may be due to these collisions of the molecules, and from the above computation it is clear that these collisions may occur as often as one million times per second, and yet each wave train will contain *five hundred million individual waves.* Such changes prevent the appearance of interference.

It should be carefully noted, however, that in light, as in sound, the system of interference bands denotes only a redistribution of the vibratory energy, and while it may be reduced to zero at some points and heaped up at others, yet the total amount of energy is the same.

488. Interference from Two Small Apertures. As already mentioned in the foregoing article interference and diffraction phenomena have much in common. In the early history of the undulatory theory it was found extremely difficult to furnish experimental proofs of interference which were free from the objection that the phenomena were confused with those of diffraction. Even the fundamental experiments of Young and Fresnel suffered from this defect. However, both on account of its historic interest and its simplicity, it seems best to begin with Young's experiment of interference from two small apertures.

Suppose a beam of monochromatic light from a narrow slit be allowed to fall upon the two small apertures A and B (Fig. 284), so that the two pencils from these points are coherent and in the same phase. Draw OM normal to AB, at its middle point, and at M erect a perpendicular PM. Now, since M is equidistant from the two apertures, it is evident that the paths traversed by the two pencils are identical and the two wave trains will reach the point M *in the same phase*, and the illumination will be a maximum. If we assume the amplitudes to be equal, then the resultant amplitude of vibration will be the sum of the two equal amplitudes and will therefore be double that of either, while the intensity of illumination will be *four times as great* as that from either source singly.

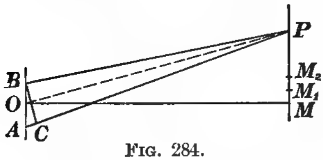


FIG. 284.

Next consider a point M_1, chosen so that the difference in path of the two pencils is $\lambda/2$ for light of the particular color used. Then $AM_1 - BM_1 = \lambda/2$, and the pencil from A will reach M_1 just *one half period behind* the pencil from B. The vibrations will be, therefore, in *opposite phase and will annul each other*, and there will be darkness at this point. If at M_2 the difference in path amount to $2\,\lambda/2$, or a whole wave length, then the vibrations from A reach M_2 *one whole period behind those from* B, and hence *coincide in phase, and the two sets of vibrations reënforce each other*, producing maximum illumination.

In general, if P be any point, either above or below M, such that

$$AP - BP = \pm\frac{n\lambda}{2}$$

then P will be a bright or dark point according as n is *even* or *odd* (Art. 121). If, therefore, the slits be illuminated by sodium light, there will appear upon the screen, normal to the plane of the paper, a series of alternately yellow and black fringes starting with M, and extending some distance both above and below this point. If either aperture be closed, the fringe system vanishes.

The distance of any band from the central band may be calculated as follows: Call the distance MP, x; let MO equal a, and denote the distance AB between the two sources by c. Draw the lines OP, BP and AP, and about P as center, with a radius BP, describe the short arc BC. AC will, therefore, represent the difference in path, $n\lambda/2$. The two angles ABC and POM are equal, since their sides are mutually perpendicular, and the two triangles ABC and OMP are similar. Also, since the angles ABC and POM are very small, their sines and tangents may be equated and we have

$$\frac{MP}{OM}=\frac{AC}{AB}$$

whence

$$x=\frac{a}{c}\cdot n\frac{\lambda}{2} \tag{559}$$

This formula shows that the distance of any fringe from the central fringe M varies directly as the wave length of the light employed. If, therefore, *white light be used*, the central fringe at M, being the position of *zero difference in phase*, *will be white*. The other parts of the system, instead of being marked by bright and dark bands, will now show a set of rainbow-colored fringes, and there will be no dark bands at all. This is because the different colors correspond to different wave lengths. Moreover, experiments shows that the fringes are *violet* on the *inner edges nearest M*, and *red* on the *outer edges*, hence we see that *violet* light has the *shortest* and *red* the *longest* wave length of the spectral colors. If with monochromatic light the quantities a, c, n and x be carefully determined, the wave length λ for the corresponding color may be computed.

***489. Fresnel's Biprism.** In order to demonstrate the interference of light as clearly as possible, and also to obviate any possible objections as to the reality of interference phenomena, Fresnel devised a number of beautiful experiments, of which the simplest is that of the biprism. An isosceles glass prism CED (Fig. 285), with the angle at E very nearly 180°, receives monochromatic light from a vertical slit at O. The two halves

of the prism CE and DE behave as two right-angled prisms put base to base, and the light from the upper half is refracted downward as if coming from a virtual source B, while that passing through the lower half of the prism gives rise to a second virtual image at A. The distance between these two virtual images A and B is smaller the smaller the angles C and D become, so that by making E very nearly 180°, the two pencils emerging from the two halves of the double prism (biprism) overlap throughout a certain region FMG upon the screen. M, being equidistant from the two virtual sources, is a bright band

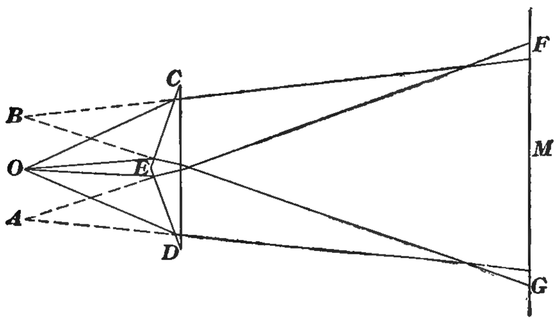


FIG. 285.

for all colors. On either side of M there appear alternately dark and bright bands, standing normally to the plane of the paper, marking the positions for which the difference of path is an *odd* or *even multiple of* $\lambda/2$ for the special wave length of light under consideration. If white light be used, the dark and bright bands are replaced by a series of rainbow-colored fringes, violet on the inner and red on the outer edges.

The distance of the nth band from the center is readily calculated. Thus let OE equal a, EM equal b, ϵ equal the angle at C or D of the prism, and μ its refractive index. The deviation δ produced by either half of the biprism may be written (Art. 451),

$$\delta = \epsilon(\mu - 1),$$

and the distance $AB = c$, becomes

$$c = 2a \sin \delta = 2a\delta = 2a(\mu - 1)\epsilon \qquad (560)$$

Therefore $$x = \frac{(a+b)}{c} \cdot n\frac{\lambda}{2} = \frac{a+b}{2\,a\,\epsilon(\mu - 1)} \cdot n\frac{\lambda}{2} \qquad (561)$$

The biprism has the advantage that the fringes are very bright and are readily obtained.

490. Interference in Thin Films. Films of any transparent substance, if sufficiently thin, exhibit brilliant colors when viewed by white light, or a series of dark and bright bands if examined by monochromatic light. Examples of such films are seen in soap bubbles, thin films of glass, thin films of oil upon water, or of oxide upon heated surfaces of polished metals. In order to comprehend the relation of these interference bands to other interference phenomena, it is necessary to conceive the two interfering pencils as produced by reflection at the upper and lower surfaces of the film. Thus, let FF' (Fig. 286) represent a uniform thin film of air inclosed between two plates of glass G and G'. When monochromatic light falls normally upon the film, a part of the light is reflected at the under side U, of the upper plate G, in the direction UB. The other part penetrates the film, and part of this light is reflected at L, the upper side of the lower plate in the same direction LD, while the remainder passes through the plate. Both reflected components of the original beam are propagated in the direction CA, and the illumination produced by UB and LD depends upon their difference of path, which *seems to be* the double thickness of the film, $2\,t$, expressed in wave lengths of the light under consideration. From our previous study of interference phenomena we should expect that, according as this difference in path $2\,t$ amounts to 0, λ, $2\,\lambda$, etc., or to $\lambda/2$, $3\,\lambda/2$, $5\,\lambda/2$, etc., that is, according as the difference in path amounts to an *even* or an *odd* number of half wave lengths, the film, to an eye looking down upon it along AC, should appear *bright* or *dark*.

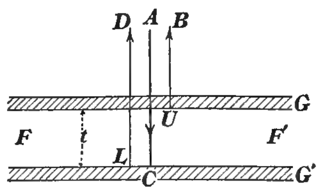


FIG. 286.

Observation shows that exactly the opposite result is attained.

When the film is made as thin as possible, *it appears black*, thus showing that for a vanishingly small thickness t the two pencils are in *opposite phase* instead of *coincidence*.

The reason for this is found in the opposite conditions of reflection experienced by the two pencils; the beam reflected at U is reflected *in* glass *against* air, while the beam reflected at L is reflected *in* air *against* glass. In our study in sound (Art. 118) we have seen that when a wave is reflected *in* a denser medium *against* a rarer one, the motion of the particles *is not reversed* (open end of pipe); while in the case of reflection *in* a rarer medium *against* a denser, the motion of the particles *is reversed* (closed end of pipe). Similar relations hold in the reflection of light, and since the reversal of the motion of the particles is equivalent to a change of phase of half a period, this relation is usually expressed by saying that in the case of reflection *in air against glass, a half wave length is lost.* Fig. 287 shows the simultaneous displacements of the ether in the two beams, where the incident wave is represented by the full line and the reflected waves by the dotted lines.

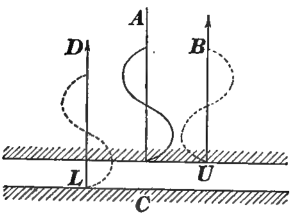


FIG. 287.

In the case of a film of vanishingly small thickness, therefore, the two wave trains are superimposed in opposite phases, and the result is darkness. The total difference of phase is therefore that due to a difference in path, $2t$, *plus* that due to difference in conditions of reflection $\lambda/2$. Hence the total retardation is $2t + \lambda/2$, and when *this quantity* amounts to an *even* number of half wave lengths, the film is bright. When it is equal to an *odd* number of half wave lengths, the film is dark.

In the case of oblique illumination it is easily shown that the retardation due to the film is $2t \cos r$, where r is the angle of refraction into the film. Hence for an eye looking down upon a horizontal film of uniform thickness, the path retardation increases with the obliquity of the light entering the eye. The film will therefore present a series of interference fringes

in the form of circles concentric about the foot of the perpendicular from the eye upon the plane of the film. The circles are the loci of points of equal phase difference.

If the film be wedge-shaped (Fig. 288) instead of plane parallel, it will be crossed, when viewed by reflected light, by a system of dark bands which run parallel to the edge of the

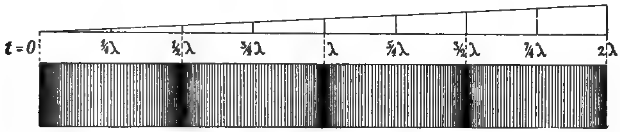


FIG. 288.

wedge, and mark the places at which the thickness of the film is $0, \lambda/2, 2\lambda/2, 3\lambda/2$ . . . etc. Since for extinction of the light the total retardation $2t + \lambda/2$ must equal an *odd number* of half wave lengths, we have *for the dark bands*

$$(2n+1)\frac{\lambda}{2} = 2t + \frac{\lambda}{2} \tag{562}$$

where n is any integer, 0, 1, 2, 3, etc., and denotes the number of the dark band under consideration. The dark band corresponding to $n = 0$ denotes optical contact of the two surfaces, or zero thickness of the film.

***491. Interferometers.** Any device whereby two pencils of light may be made to interfere, and the resulting phenomena studied, is, properly speaking, an interferometer. Certain forms of the instrument present peculiar advantages, and merit special description. In the interferometer originally devised by Fizeau, and modified and improved by Abbé and Pulfrich, the two pencils are made to interfere after normal reflection upon the surfaces of two horizontal plates of glass inclosing a film of air (Fig. 289) in which interference occurs exactly as described in the preceding article. The other parts of the instrument are for convenience of observing and measuring the fringe system.

The under plate of glass B is actuated by a micrometer screw S, so that the thickness t of the air film may be varied at will.

The plates are slightly inclined to each other, thus giving the air film a wedge shape. The interference bands are then straight lines parallel to the edge of the wedge (Fig. 288). On decreasing the thickness of the air film the bands move toward the thicker part of the wedge; on increasing the thickness they move in the reverse direction. Hence to measure the wave length of any colored light, the system is illuminated with that particular light and the micrometer head turned until a definite number b of dark bands have passed over a certain point, usually a small reference circle on the under side of the upper plate A. The amount s, by which the thickness of the film has been changed, is determined from the micrometer head and the quantities b and s inserted in the general formula.

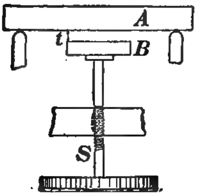


FIG. 289.

Thus for the dark band under the circle we have by (562)

$$(2n+1)\frac{\lambda}{2} = 2t + \frac{\lambda}{2} \text{ (before)}$$

also since to pass from any dark band to the next higher one we must introduce a difference of path of *one whole wave length*, we have, after passing over b dark bands,

$$[(2n+1)+2b]\frac{\lambda}{2} = 2(t+s) + \frac{\lambda}{2} \text{ (after)} \qquad (563)$$

Whence
$$2b\frac{\lambda}{2} = 2s$$

or
$$\lambda = \frac{2s}{b} \qquad (564)$$

***492. The Michelson Interferometer.** In the Michelson interferometer, a beam of light from the source Q (Fig. 290) falls at an angle of 45° upon the half-silvered face of a plane parallel glass plate A, where it is divided into two pencils, one of which is transmitted and passes to the mirror D, the other is reflected to the mirror C. These two mirrors are set so as to return the two pencils upon their paths to the point A, where the first

is reflected and the second transmitted to E. A second plane parallel plate B, of identically the same thickness as A, is inserted in the path of the reflected ray to make the paths traversed by the two pencils meeting at E equal, in case D and C are symmetrically placed with respect to A.

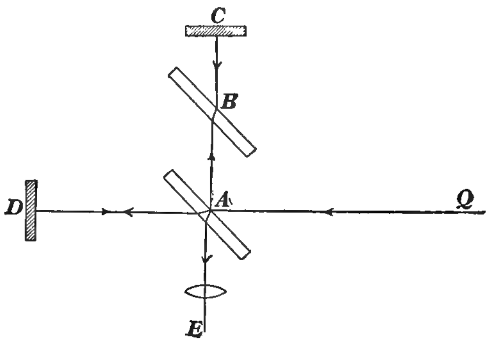


FIG. 290.

Now the transmitted ray AD has passed through the plate A three times and has been reflected once *in glass against air*. The reflected ray AC has passed twice through B, once through A, and has been reflected once *in air against glass*. When the two pencils have traversed equivalent paths, they are in condition to interfere, owing to the half wave-length difference in phase introduced by opposite conditions of reflection at A. One of the mirrors D is movable in the direction AD by means of a micrometer screw. Wave lengths may be measured as in the Pulfrich-Abbé instrument.

The Michelson interferometer has the advantage of a wide separation of the two pencils of light, and it may consequently be applied to an almost endless variety of physical problems.

CHAPTER LVII

DIFFRACTION

493. Diffraction through a Narrow Slit. If a strong beam of parallel rays be passed through a narrow vertical slit into a darkened room and received upon a white screen some two or three meters distant, there will be seen a central band of white light, broader than the dimensions of the slit would justify from strictly rectilinear propagation, and on either side a series of colored fringes. It is evident that through a narrow slit the light does not travel in straight lines even approximately, but bends around the edges, and spreads out in all directions, from all points of the opening as new centers of subsidiary waves. This phenomenon is called *diffraction*, and the fringes are called *diffraction fringes*. As has already been pointed out, diffraction is a species of interference between waves arising from different points on the *same wave front*.

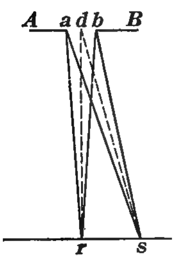


FIG. 291.

Let ab (Fig. 291) represent a horizontal section through the vertical slit, and let r be a point on the screen, such that ar is equal to br. Then the disturbances from all points in the slit will reach r in practically the same phase. They will therefore reënforce each other, and r will be a bright point for all colors, and the central band will be *white*. If now a point s be taken near r, such that

$$as - bs = \lambda \tag{565}$$

for some definite color, as the violet, then the waves setting out from b would reach s one whole period ahead of the waves from a. If d be a point midway between a and b, then a wave from

d would arrive at s just one half period later than the one from b, and hence these waves would annul each other. In like manner every elementary wave from a point in da would be annulled by a wave from a corresponding point in bd. The result at s would therefore be zero *for violet light.* Consequently s marks the extreme edge of the central illumination from the slit ab, for violet light, or the broad image of the slit in violet light extends on either side of r, through the distance rs.

If, on the other hand, a point s_1 be chosen, such that

$$as_1 - bs_1 = 3\frac{\lambda}{2} \tag{566}$$

for violet light, then the slit ab can be divided into three parts, from *two of which the waves will interfere as before*, while the remaining third will produce violet illumination.

In general, if

$$as - bs = \pm n\frac{\lambda}{2} \tag{567}$$

for points on either side of r, outside the central band, we shall have a series of bright and dark bands, for monochromatic light. The bright bands will correspond to the points where n is *odd* and the dark bands to the points where n is *even.* For white light the fringe system is a series of rainbow-colored bands, each band being violet on the inner, and red on the outer edge, thus showing again that violet has the shortest, and red the longest wave length of the spectral colors.

494. The Diffraction Grating. If, instead of a single slit, a series of parallel, equidistant slits be ruled upon a piece of smoked glass, or better upon the opaque film of a photographic plate, then the colors of the diffraction fringes are much more lively, and the phenomena in strong sunlight are very beautiful. Such a ruled surface is called a *diffraction grating*, and the resulting spectra are called *diffraction* spectra.

Let M_1N_1, N_1M_2 (Fig. 292) represent the transparent and opaque parts respectively of a diffraction grating, and suppose the light from a collimator slit to come in a parallel beam

from the left striking the grating normally. Then along the direction indicated by the dotted line from M_2 we shall have a bright band, which may be focused by a lens into an image of the slit.

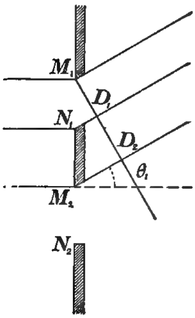


FIG. 292.

Next suppose that for light of wave length λ, emerging in direction N_1D_1, M_2D_2, the difference in path of the waves from corresponding points in adjacent slits of the grating as M_1 and M_2 should be one wave length, or $M_2D_2 = \lambda$; then it is clear that the light from the first slit will be in advance of the light from the second slit by *a whole wave length*, and ahead of that from the third slit by *two wave lengths* and so on. The light from all the slits will, therefore, agree in phase, and the resultant illumination may be focused by the lens into a *diffracted* image of the slit. This is called the *first diffracted image* of the slit, or a diffraction image of the *first order.*

If the direction of the light for the first spectrum make an angle θ_1 with the normal to the grating, and M_1N_1 be set equal to a, and N_1M_2 equal to b, then since the angle $D_2M_1M_2$ is equal to θ_1, we have

$$M_2D_2 = \lambda = (a + b) \sin \theta_1 \tag{568}$$

Similarly for an angle θ_2 such that the difference in path between corresponding points of adjacent slits is 2λ, we have for the second spectrum

$$2\lambda = (a + b) \sin \theta_2$$

for the third spectrum

$$3\lambda = (a + b) \sin \theta_3$$

and so on. From equation (568) we have

$$\sin \theta_1 = \frac{\lambda}{a + b} \tag{569}$$

which shows that the sine of the angle of diffraction is directly proportional to the wave length, hence in a diffraction spectrum

the violet light is *deviated least* and the red light *most.* Moreover, since the deviation of every color is directly proportional to the wave length, it follows that each color lies in its proper place and that in such a spectrum there can be no irrationality of dispersion.

Again, since $\sin\theta_1$ varies inversely as the *grating constant* $(a+b)$, it is clear that the lengths of diffraction spectra from two different gratings are inversely proportional to the constants of the two gratings and hence may be compared at once (Art. 467).

It is also to be noted that when λ becomes equal to $a+b$, the value of $\sin\theta_1$ becomes unity. This means that for waves of this length the angle of diffraction becomes 90°. Hence for measurements in the *infra red* end of the spectrum, gratings of large constant must be used.

Rutherford ruled gratings upon glass having 700 lines to the millimeter. The magnificent gratings of Professor Rowland are ruled upon speculum metal and have in some cases as many as 1700 lines to the millimeter. In these gratings the spectra are formed by light *reflected* from the ruled surface rather than transmitted by it. Glass gratings give reflected as well as transmitted spectra. The formula[1] for the reflection grating is slightly more complex than for the case here considered.

495. Measurement of Wave Lengths. The diffraction grating affords the simplest method for the measurement of wave lengths of light since, if the grating constant $(a+b)$ be known, the process is reduced to the measurement of a single angle, which can be made with great accuracy. A grating is mounted vertically upon the table of a spectrometer and so adjusted that the beam of parallel rays from the collimator strikes the grating surface normally. The slit is illuminated with monochromatic light, for example that furnished by a Bunsen burner carrying a tip of asbestos paper saturated with sodic nitrate. On turning the telescope so as to look directly into the collimator, the direct, or central, image of the slit should be seen sharply focused, when the telescope and collimator are set for parallel rays. On either side of this central image will be seen the yellow diffracted

[1] *For a discussion of the more general case, see Manual, under Diffraction.*

images of the slit, of the first, second, third and fourth order. Owing to the small fraction of the aperture in each slit which is effective in producing illumination, the intensity of the diffracted images falls off very rapidly. The vertical cross hair of the telescope is set upon the first diffracted image on each side of the central image, and the readings are taken. One half the difference between these readings is θ_1. In the same way the values of θ_2, θ_3 and θ_4 are determined. Then

$$\lambda = (a + b) \sin \theta_1 = (a + b) \frac{\sin \theta_2}{2} = (a + b) \frac{\sin \theta_3}{3}, \text{ etc.}$$

Conversely, if light of a known wave-length be used, the grating constant $(a + b)$ may be at once determined.[1]

496. Bright Line Spectra. Through the investigations of Kirchhoff and Bunsen, 1856–60, the following fundamental facts were established concerning the three more important types of spectra.

When light from an incandescent gas or vapor is examined by means of a prism, its spectrum is seen to consist of a number of bright lines, colored images of the slit, which are always the same for the same gas under the *same conditions of temperature and pressure.* Thus the spectrum of sodium vapor at the temperature of the Bunsen burner consists of a single pair of bright yellow lines corresponding to the Fraunhofer lines D_1 and D_2. The spectrum of lithium at this temperature consists of a single line in the deep red. The light from hydrogen in a Geissler tube shows four well-marked lines, one in the red and one in the blue corresponding to the Fraunhofer lines C and F, and two fainter lines in the violet.

Such a spectrum is called a *bright line spectrum*, and its presence indicates to us that the source of light is a mass of incandescent gas or vapor under a pressure so low as to allow the gas molecules sufficient freedom of motion to execute whatever form of vibration they will. From the fact that the spectrum of a chemically pure substance in the gaseous form may contain *numerous bright lines*, we are driven to the conclusion that

[1] *For experimental determination of wave lengths of light, see Manual, Exercise 90.*

the molecule of such a gas may execute a number of different vibrations at the same time, just as a string or a plate may vibrate in a number of different modes, and produce a number of corresponding tones at the same time. Under this aspect of the case the characteristic bright lines in the spectrum of any gas at a given temperature may be regarded as representative of the *free vibrations which its molecules can execute at that temperature.*

497. Continuous Spectra. When the light from an incandescent solid or liquid, or from a mass of incandescent gas *under high pressure*, is analyzed, the spectrum is found to contain all colors from red to violet, and to show no discontinuities at any point. Such a spectrum is called a *continuous spectrum* and shows that the source is a mass of incandescent solid, liquid or gas at high pressure. The spectra from molten metals, from the filaments of incandescent lamps or from the carbon tips of an arc lamp are all continuous spectra.

In the case of such a luminous source, it is clear that the molecular motions due to incessant molecular collisions must be extremely irregular and constantly interrupted. The molecules have practically no mean free path, and no time in which to execute their characteristic vibrations. The result is a confused, jangled mass of vibration of every possible frequency, which the eye interprets as light of all colors, *i.e. a continuous spectrum.*

498. Dark Line, or Absorption Spectra. Absorption spectra are produced when light from an incandescent solid, liquid, or gas at high pressure is passed through a layer of some unequally transparent medium, and then analyzed. The spectrum is seen to be crossed by one or more dark lines or bands, indicating that in these regions the energy of the spectrum has been absorbed by the medium under investigation. Liquids are examined for absorption by placing them in tanks with parallel sides of plane parallel glass plates.

Many substances present characteristic absorption spectra. Thus a piece of cobalt glass absorbs all colors except a small strip in the red, and in the blue end of the spectrum. The absorption spectrum of chlorophyl shows a dense black line in the

red, while blood, even in very dilute solution, shows two characteristic bands in the green.

In the case of a gas the absorption spectrum exhibits one or more dark lines sharply defined upon the continuous spectrum of the source. These lines are characteristic for the gas and correspond to certain bright lines emitted by the same gas when raised to incandescence. Thus at the temperature of the Bunsen burner, sodium vapor absorbs only the yellow rays belonging to the D lines.

The principle of absorption is merely the principle of resonance (Art. 131) applied to the motion of ether particles. The light emitted by a vibrating atom of a heated gas may be considered as representative of the vibrations which it can execute. If those same vibrations fall upon the gaseous atoms, they will covibrate or take up the vibratory motion, just as a tuning fork will respond to vibrations of its own frequency but to no others.

It is to be noted further that the glowing gas acts at the same time both as an absorbing and as an emitting layer. If light from a source at a temperature higher than that of the gas pass through the gas, then the gas molecules take up more energy than they give out, or light is absorbed by the gas. But if the gas be at the higher temperature, then the gas molecules give out more energy than they absorb, and light of that particular wave length is added to the light of the source. In the first case the lines are darker the greater the degree of absorption, *i.e.* the greater the difference of temperature. In the second case the line shows as a bright line on the continuous spectrum. For equality of temperature between the sources the line vanishes.

499. Spectrum Analysis. Since the character of the light emitted by an incandescent gas depends first of all upon the vibrations of its constituent atoms, it follows that a study of the light emitted by a glowing gas gives us direct testimony concerning its chemical composition. Hence, if the bright line spectrum of any substance be once definitely known, then whenever this spectrum presents itself we may conclude at once that the given substance is present in the source of light,

whether that source be a Geissler tube in the laboratory or a fixed star in the depths of space. This is the method of spectrum analysis. A minute quantity of a salt is introduced into the colorless flame of a Bunsen burner and the light examined by the spectroscope. The method is most useful in the detection of the metallic constituents of salts, since at the temperatures necessary to vaporize a salt and tinge the flame, the spectrum is generally independent of the acid constituent.

The method is characterized by its ease and rapidity, and especially by its exceeding sensitiveness. In the flame of the Bunsen burner, 1/14,000,000 of a milligram of sodium is sufficient to show the characteristic sodium lines, while in the spark of an induction coil, 1/80,000,000 of a milligram of lithium may be detected. On account of the extreme sensibility of the method it has led to the discovery of numerous new elements, which have been present in minute quantities as impurities in the substances under examination, and have revealed themselves through characteristic new lines in the spectrum. Among the elements so discovered may be mentioned caesium, rubidium, thallium, indium and gallium.

Spectrum analysis gives at once the explanation of the Fraunhofer lines in the solar spectrum, and enables us through comparison with bright line spectra from known sources to prove the presence of many chemical elements in the sun. Thus the two D lines of the solar spectrum coincide exactly with the two yellow lines of the spectrum of sodium, and Kirchhoff concluded that there must be sodium vapor in the sun's atmosphere. By means of the Fraunhofer lines Rowland has definitely proven the presence of thirty-six chemical elements in the sun.

500. Peculiarities of Spectra. Again, it should be noted that while the spectrum of any substance is characteristic of that substance, and furnishes a reliable criterion for conclusions concerning the constitution of its molecule, yet the same substance may exhibit different spectra for different conditions of pressure and temperature. It seems reasonable to suppose that a complex molecule is capable of more varied forms of

vibration than a simpler one, and hence it seems likely that a complicated spectrum corresponds to a complicated molecular structure, and a simple spectrum to a simple molecular structure. Experiment seems to show that *each compound that can exist at the temperature at which light is emitted has its own spectrum.* As the temperature of a solid rises, the spectrum changes correspondingly.

From the first appearance of color, the spectrum grows to the completed, *continuous* spectrum. For a slightly higher temperature, but one at which the compound molecule can still exist, the spectrum is marked by the appearance of bright parts, which are not yet sharp lines, but rather broad bands, set off by bright lines which shade off into darkness on one side. Such a spectrum is called a *band spectrum* or a *fluted spectrum,* since it has the fluted appearance of a Greek column. Such spectra correspond to relatively low temperatures and complicated molecular structure, usually that of a chemical compound.

For still higher temperatures the compound molecule breaks up into its constituent atoms, and the *bright line* spectrum appears. This, as we have seen, corresponds to a highly heated gas under low pressure.

Again, the spectrum of an element may contain but a few bright lines which seem to be arranged in some apparently definite, rhythmical order, as in the spectrum of lithium, or it may contain a thousand lines arranged in apparently the greatest confusion as in the case of iron. These lines may also occur singly as in lithium, or in pairs as in sodium, or in triplets as in magnesium. These groups of single lines, or pairs, or triplets recur at regular intervals, the intervals growing shorter as we approach the violet end of the spectrum. The analogy between these rhythmically recurring groups of lines and the overtones produced by a sounding body is certainly very striking. Through the investigations of Kayser and Runge, certain harmonic relations between the vibration frequencies of the spectral lines of many of the elements have been established, but these relations are by no means simple, nor have they as yet been shown to exist in the case of all the elements.

CHAPTER LVIII

RESOLVING POWER OF OPTICAL INSTRUMENTS

501. Resolving Power of the Telescope. The performance of every form of optical instrument reaches a limit imposed by the nature of light itself. If a plate ruled with fine parallel equidistant lines be examined, either by the unaided eye, or by means of a telescope or a microscope, there will in each case be found a limiting distance s, between the lines, below which they are no longer seen as separate and distinct lines. The limiting angle subtended by s is termed *the limit of the resolving power* of the instrument in question, or more briefly, *the resolving power*. The principles involved in the determination of this limiting value may be best explained from the telescope.

Suppose ES (Fig. 293) to be a narrow slit and O its middle point. Let p be any point on a line normal to ES through O,

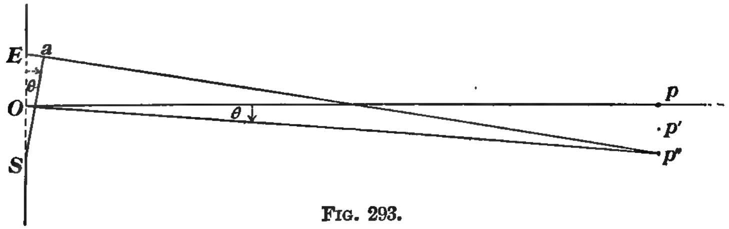


FIG. 293.

and at a distance from the slit. Then on directing the slit toward a small brilliant source of light, p is a bright point for light of any color entering the slit. Next let p' be a point on a perpendicular to Op through p, such that $Ep' - Sp'$ is equal to $\lambda/2$ for light of some definite wave length. Then waves starting from the points E and S will reach p' in opposite

phase, and so annul each other. But it is *only for the extreme edges* E and S that this is true, hence there will be *some* illumination at p'. But if we take a point p'', such that $Ep'' - Sp''$ is equal to λ, then p'' is a dark point for light of wave length λ (Art. 493). We should therefore have a series of bright and dark bands above and below the central broad bright band whose center is at p. These bands lie parallel to ES and normal to the plane of the paper, and the point p'' marks the *first dark band* on the lower side of p.

If now ES be regarded as the diameter D of a telescope objective, we have a series of bright and dark rings concentric about the bright central disk, whose center is at p. This bright disk may be regarded as the image of a point source of light of wave length λ. Calculation shows the radius of the first dark ring to be slightly larger than the corresponding value derived for the first dark band. Thus the radius of the first dark ring is $1.2\, pp''$, or the diameter is $2.4\, pp''$.

Let ES equal D, Op equal F, the focal length of the objective, d equal $2.4\, pp''$, the linear diameter of the image of the point source or artificial star, and Ea equal λ. About p'' as a center, with $p''S$ as radius, describe the short arc Sa. Then the angles ESa and pOp'' are equal, and equating values of sine and tangent we have

$$\frac{Ea}{ES} = \frac{pp''}{Op} \tag{570}$$

or

$$\frac{\lambda}{pp''} = \frac{D}{F} \tag{571}$$

whence

$$2.4\, pp'' = \frac{2.4\, \lambda\, F}{D} = d \tag{572}$$

where d is the linear diameter of the image of the artificial star. The angle subtended by any image at C, the center of the objective, is d/F, which (Fig. 294) is readily seen to be the same as that subtended by the object at the same point. Then $d/F = 2.4\, \lambda/D$, or *the angular diameter* of the star image.

Hence if two star images are to be seen as separate disks, then the angular separation of their centers must be at least

$$\frac{d}{F}=\frac{2.4\,\lambda}{D} \tag{573}$$

If we take λ as 0.00056 mm, and D as one inch or 2.54 cm, then the angular diameter of the star disk is $10''.92$; or the images of two stars which subtend an angle $10''.92$ would appear in the telescope with an objective one inch in diameter, with their disks just touching, if the light could be traced out to the edges of the diffraction disks. On account of the faint light of the stars, however, the extreme edges of the disks are invisible, and under most favorable circumstances two stars can be separated, which are a little less than half this distance apart, or *the limit of the resolving power of a telescope whose objective is* 1 *inch in diameter is about* $5''$. Hence to find the resolving power of any telescope divide $5''$ by the diameter of the objective in inches.

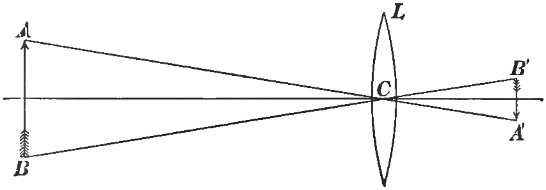


FIG. 294.

502. Resolving Power of the Eye. It is shown in the last article that the resolving power of a telescope depends simply upon the wave length of light employed, and upon the diameter of the objective, or two objects to be resolved must subtend an angle

$$\phi>\frac{2.4\,\lambda}{D} \tag{574}$$

In the case of the eye, the crystalline lens has a refractive index of 1.4, and hence the wave length λ in air becomes $\lambda/1.4$ in the lens of the eye. The diameter of the pupil corresponds to the diameter of the objective of the telescope, and putting this diameter equal to 4 mm, we have as the resolving power of the eye

$$\phi=\frac{2.4}{4}\cdot\frac{\lambda}{1.4}=49.''5 \tag{575}$$

The actual limit is *about one minute*. This means that a normal eye can see two lines separated, whose distance apart subtends at the eye *an angle of one minute of arc*.

It is interesting to note that the "rods and cones," or light sensitive elements of the eye, subtend the same angle, one minute, at the nodal point of the eye. From this it appears that whether we consider the resolving power of the eye as depending upon the smallest distance between two nerve endings capable of receiving separate stimuli, or whether the eye be regarded as a simple lens, the theoretical resolving power comes out the same in either case. The resolving power of the eye may be readily determined by the following experiment.

Draw a series of equidistant lines upon a piece of white paper, making the lines and spaces of about the same width, and determine the distance from the eye at which the lines under bright light can just be seen resolved. The angle subtended at the eye by the distance between two lines will give the resolving power for the eye in question.

***503. Resolving Power of the Microscope.** It has been pointed out that the magnification of a simple lens soon reaches a limit owing to the short working distance and the minute dimensions of the lens. A more serious difficulty is found in the fact that as the lens grows smaller the dimensions of the opening are no longer very great as compared to the wave length of light. The image of a point source is therefore a diffraction disk of definite radius, and this radius increases as the diameter of the lens decreases. It can be shown experimentally that for a lens whose diameter is less than one one thirtieth of an inch the confusion arising from the increased diffraction is very great.

The *angular aperture* 2α is the angle included between the extreme rays, which can pass through the microscope objective from a point distinctly seen on the axis of the instrument. If μ be the index of the medium *from which* the light enters the objective, then

$$\mu \sin \alpha = N \tag{576}$$

is called the *numerical aperture* of the objective. In air μ is 1 and the maximum value for N is also 1. Since this would

denote an angular aperture of 180°, it is obvious that the greatest diameter D that can be used in a simple lens is twice the focal length F, or

$$N=\frac{D}{2F} \tag{577}$$

where N has its maximum value of 1.

By placing upon the cover glass a drop of some fluid in which the objective may be immersed, the factor μ may be varied at pleasure and the aperture N may be increased to 1.4 or even more. Such a lens system is termed an *immersion system.* If a liquid be used whose index is the same as that of the objective, it is termed *homogeneous immersion.* If a grating of constant $(a+b)$ or d be viewed by a microscope in direct light, then the diffraction pattern in the image will resemble the structure of the object (grating), only when all the rays diffracted by the object of sufficient intensity to produce appreciable effects in the focal plane of the objective are received by the objective. Hence the resolving power depends upon the numerical aperture of the objective. When the grating is viewed by direct light, therefore, the first maximum from the center of the field lies in the direction $\sin\theta_1$ or λ/d. Hence if the grating is to be seen resolved, that is, if the lines are to be seen separated in the image, the objective must receive rays whose inclination is at least equal to θ_1 where

$$\sin\theta_1=\frac{\lambda}{d} \tag{578}$$

In an immersion system, the wave length λ' in the fluid is equal to λ/μ, where μ is the refractive index of the immersion fluid, and λ is the wave length in air. In this case

$$\mu\sin\theta_1=\frac{\lambda}{d} \tag{579}$$

for resolving a grating of constant d. But since $\mu\sin\alpha$ is the numerical aperture N, of the objective, then to resolve the grating, N or $\mu\sin\alpha$ must equal $\mu\sin\theta_1$, or

$$N=\frac{\lambda}{d}$$

Hence the smallest distance d which can be resolved by a microscope of numerical aperture N, in direct illumination, is

$$d = \frac{\lambda}{N} \tag{580}$$

In the case of oblique illumination this may be reduced one half under *the most favorable conditions*, or

$$d = \frac{\lambda}{2N} \tag{581}$$

Since the resolving power *increases as d decreases, it is clear that the resolving power af the microscope varies inversely as the wave length of light used, and directly as the numerical aperture of the objective employed.*

Taking λ as 1/50,000 of an inch or 0.000508 mm, for greenish blue light, and N as 1, we have for *oblique illumination, in air, under most favorable conditions,*

$$d = \frac{\lambda}{2N} = 0.00001 \text{ inch or } 0.000254 \text{ mm}$$

This means that under the above conditions a grating having 100,000 lines to the inch could be seen resolved. In the case of homogeneous immersion this limit may be extended to 135,000.

***504. Resolving Power of a Grating.** If light fall upon a diffraction grating of constant $(a + b)$ or d, at an angle of incidence α, and the transmitted light be diffracted at an angle β, then it may be shown that for the points of maximum illumination in the mth spectrum the maximum phase difference between wave systems from corresponding points of adjacent slits of the grating is

$$d(\sin\alpha + \sin\beta) = m\lambda \tag{582}$$

This form of the equation for a grating applies equally well to reflection or transmission gratings provided α and β lie on the same side of the normal to the plane of the grating.

In order to investigate the resolving power of the grating it is necessary to inquire into the conditions necessary to separate

the mth spectral image of wave length λ from the mth image of wave length $\lambda + d\lambda$, where $d\lambda$ is a very small fraction of a wave length. The solution given by Lord Rayleigh is as follows: In a grating of n lines, the mth spectrum lies in such a direction that the phase difference between corresponding points of adjacent slits is $m\lambda$, and between wave systems from the extreme slits it is $mn\lambda$. The nearest points of minimum intensity on either side of the mth maximum correspond to phase differences between the extreme slits of $mn\lambda \pm \lambda$. If now the mth maximum for wave length $\lambda + d\lambda$ fall in the position of minimum intensity for wave length λ, that is for a phase difference $mn\lambda + \lambda$, then the two lines are seen sharply separated, and we may equate length of paths for this point, or

$$(mn + 1)\lambda = mn(\lambda + d\lambda) \tag{583}$$

whence

$$\frac{\lambda}{d\lambda} = mn = r \tag{584}$$

the resolving power of the grating. The quantity $r = \lambda/d\lambda$ indicates the reciprocal of the fraction of a wave length by which two lines must differ in order to be completely separated. For example, the two sodium lines D_1 and D_2, having wave lengths 5896 and 5890×10^{-7} mm, differ by 6 units, or the ratio $\lambda/d\lambda$ may be said to have, in round numbers, the value 6000/6, or 1000. A grating which will separate the D lines must therefore have a resolving power of at least 1000. Also, since r is $\lambda/d\lambda$ or mn, it is plain that if the sodium lines are to be seen separated in the *first* spectrum the grating must have at least 1000 lines, while for the second spectrum 500 would suffice.

Again, if the equation for the grating be multiplied by n, the number of lines on the grating, we have

$$mn\lambda = nd(\sin\alpha + \sin\beta) \tag{585}$$

But mn is equal to r, the resolving power of the grating, and nd is equal to b, the breadth of the ruled surface of the grating, hence we may write

$$r = \frac{b}{\lambda}(\sin\alpha + \sin\beta) \tag{586}$$

This equation shows that the resolving power is a maximum for α and β each equal to 90°, or

$$r_{\max} = \frac{2\,b}{\lambda} \tag{587}$$

This value, however, can never be attained, since it corresponds to an infinitely small bundle of rays. If either α or β be made zero, the other angle may amount to 60°. By the use of the Abbé-Littrow autocollimation principle (Art. 452), in which telescope and collimator are combined, it is possible to make α equal to β equal to 45° to 50°; so that practically the maximum resolving power may be set down as between

$$r = \frac{7\,b}{5\,\lambda} \text{ and } r = \frac{3\,b}{2\,\lambda} \tag{588}$$

From this it follows that for the largest Rowland gratings in which b is 13.2 cm, the resolving power for λ equal to 5500×10^{-7} mm is about 375,000, if the Abbé-Littrow method be adopted. Other arrangements give much less. In the usual arrangement of the Rowland concave grating the resolving power probably does not exceed 100,000.

Michelson has recently succeeded in ruling gratings having a ruled surface of ten inches. These gratings when used in the position of autocollimation would therefore, under the best conditions, give for λ equal to 5500×10^{-7} mm a resolving power of about 680,000. It is important to note that the resolving power of a grating varies inversely as the wave length and consequently is approximately twice as great in the violet end of the spectrum as in the red.

CHAPTER LIX

POLARIZATION

505. Polarization of Light. Throughout the various optical phenomena thus far studied, there has been no indication as to the nature of the ether vibrations which have been assigned as their cause. If, as in sound, the direction of vibration be *in the line of propagation*, in other words, if the vibration be longitudinal, then there will be nothing to distinguish the beam of light when viewed from one side, from its aspect when viewed from another. If a guitar string be plucked or a violin string be bowed so as to cause it to vibrate horizontally, then the entire excursions of its vibrating parts are confined to that horizontal plane and the motion is linear and simple harmonic in that plane. A card having a narrow slit cut in it slightly wider than the diameter of the string and several centimeters long may be passed over the string *with the slit horizontal*, without disturbing the motion of the string in any way. If, however, the card be rotated in its own plane through 90°, so that the slit may stand vertical, the horizontal vibration is at once extinguished, although a vertical vibration would now be rendered possible. It is clear that the card in any position would have no influence upon the *longitudinal vibration* of the string.

The transverse vibrations of a string are therefore such as to enable us to distinguish *its sides*, or to give to the string a *two-sidedness* or *polarity*. If now it could be shown that a beam of light behaves in a similar way, it would indicate the presence of transverse vibrations. If a beam of ordinary light be allowed to fall normally upon a plate of tourmaline cut parallel to the axis of the crystal, the light which emerges will be found to possess the

two-sided property of the vibrating string. If we allow the light from one plate of tourmaline to fall normally upon a second similar plate, we shall find that it passes freely through the second when the two plates are parallel as at AB (Fig. 295). If, however, the second plate be rotated about the beam as axis, as in $A'B'$, the emergent light gradually diminishes in intensity and is entirely extinguished when the two plates stand at right angles to each other, as in $A''B''$. If the second plate be rotated still further, the light again increases; and when the plates are again parallel, it reaches its full intensity, to be extinguished again when the plates stand at right angles to each other. It thus appears that when the plates are crossed, the light from the first plate is stopped by the second, just as the vibrations of the string were stopped by the slit. The light has thus been changed in its nature so as to exhibit a *two-sidedness* or *polarity* in one plane and is therefore said to be *plane polarized.* We also conclude that the displacements in the ether are *transverse* to the line of propagation of the light.

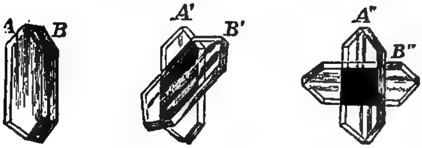


Fig. 295.

Again, since the light emerges freely from the first plate of tourmaline, no matter how it be rotated about the beam as an axis, we conclude that in ordinary light the transverse vibrations occur in all possible planes through the axis of propagation (Fig. 296 *a*); whereas, the light transmitted by the first plate is due to the vibrations parallel to the longer axis of the plate (Fig. 296 *b*). For this reason, ordinary light is considered to be made up of a mixture of light polarized in all possible planes, due to the continuous change of the plane of polarization about the line of propagation as an axis.

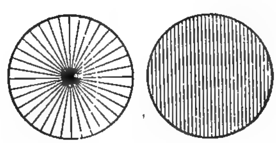


Fig. 296.

506. Polarization by Reflection. If a plate of unsilvered glass A (Fig. 297 *a*), blackened upon its rear surface, be placed

in a beam of ordinary light at an angle of incidence of about 57°, the light reflected from such a mirror will be found to be *plane polarized.* This may be demonstrated by testing the beam by means of a plate of tourmaline, or by receiving the reflected beam upon a second similar mirror *B* (Fig. 297 *a*), whose plane of incidence coincides with that of the first. The light in this position is freely reflected from the mirror *B*. If now the mirror *B* be rotated about the beam *AB* as an axis, the light reflected from the second mirror will gradually diminish in intensity, until when the planes of incidence of the two mirrors are at right angles to each other, it vanishes entirely, *A′B′* (Fig. 297 *b*), as in the case of the tourmaline plates, and reappears again in its original intensity when the mirror is rotated through another 90°.

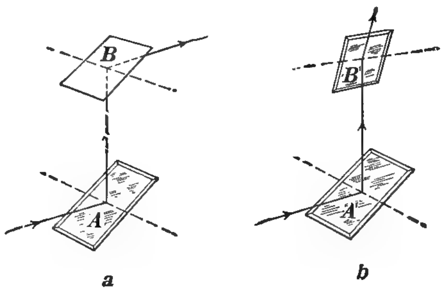


FIG. 297.

It should be noted that in the case of the crossed mirrors, the light reflected from *B′* is not zero except for a particular angle of incidence for both mirrors. This angle is called the *angle of polarization* and for glass it is between 55° and 57°. Other transparent substances may be used as mirrors, and for each substance there has been found an angle of incidence, depending upon the substance, which gives a *maximum* of polarization.

If we suppose the beam *AB* to be of unknown origin, then it may be *analyzed*, that is, its condition of polarization may be examined by means of the second mirror. If, on rotating the mirror *B* about the beam as axis, the reflected light show no change in intensity, we conclude the beam *AB* is one of ordinary light. If, however, for certain positions of the mirror *B*, the light vanish, it is *plane polarized*, and the *plane of incidence in which it is reflected most copiously from the second mirror is called*

the plane of polarization. According to the theory of Fresnel, the vibrations of plane polarized light are perpendicular to the plane of polarization. Thus the direction of the vibration in light polarized by reflection is normal to the plane of incidence, that is, *it is parallel to the surface of the mirror.*

*** 507. Brewster's Law.** It has already been stated that the angle of polarization differs for different substances. In 1811 Sir David Brewster discovered the remarkable fact, that when light falls upon a transparent substance, at the polarizing angle, the reflected and refracted beams *are at right angles to each other.* Thus, if IO (Fig. 298) be the incident ray, OR the reflected ray, and OB the refracted ray, then by Brewster's law the angle $BOR = 90°$. This law may be put in another form. Since the angle between the reflected and the refracted rays is equal to 90°, then

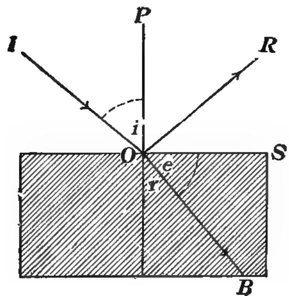


FIG. 298.

$$i + r = 90° \text{ or } \cos i = \cos(90° - r) = \sin r, \qquad (589)$$

therefore

$$\frac{\sin i}{\sin r} = \frac{\sin i}{\cos i} = \mu \text{ or } \mu = \tan i \qquad (590)$$

Now since μ is greater than unity, we learn from Brewster's law that the polarizing angle is always greater than 45°; and further, that if the index of refraction for a transparent substance be known, the polarizing angle can at once be deduced. Brewster's law has been verified by Seebeck for a number of refractive media.

508. Polarization by Refraction. If a beam of ordinary light fall upon a thin transparent glass plate at an angle of about 55° to 57°, a part of this light is reflected, and by this reflection polarized, the plane of polarization being *in* the plane of incidence. The other part of the light is transmitted, and if examined will be found to show traces of polarization in a plane *at right angles* to the plane of incidence. If the light

emerging from the first plate be passed through a second parallel and similar plate, the amount of polarized light in the emergent beam is increased, and after passing through some eight or ten such parallel plates the transmitted light is found to be completely polarized. Such an arrangement is called a *pile of plates* and may be used either as a polarizer or as an analyzer in optical apparatus.

The pile of plates may also be used to replace one or both the mirrors shown in Fig. 297 *a*. When the plates are used to replace mirror *B* (Fig. 297 *a*), and the planes of mirror *A* and the plates are made parallel, the light reflected from *A* is also reflected from the first plate of the pile as it would have been from mirror *B*. If, however, the plates be rotated into the position of *B'* (Fig. 297 *b*), the light is no longer reflected by the plates *but is wholly transmitted.* Finally, if two piles of plates be used, they behave toward each other exactly as two mirrors.

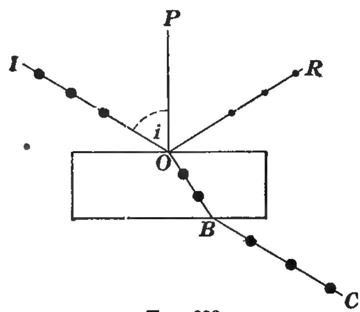


FIG. 299.

The action of the plates of glass may be understood from the following considerations: Let a beam of light fall upon a glass plate at the polarizing angle. If the light be already polarized in the plane of incidence (Fig. 299), then about $\frac{1}{7}$ of the incident light is reflected along *OR*, the rest penetrates the plate. The vibrations of the ether particles parallel to the surface of the plate are represented by the dots upon the path of the ray. If the incident light be polarized at right angles to the plane of incidence (Fig. 300), then *the entire beam*

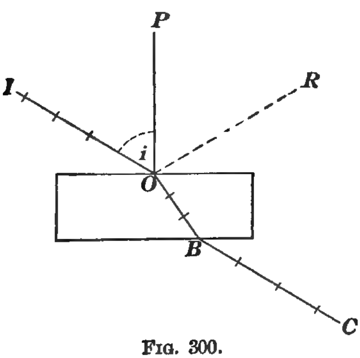


FIG. 300.

penetrates the plate and no light is reflected along *OR*. In this case the vibrations in the plane of incidence are indicated as shown in Fig. 300.

509. Double Refraction. When a ray of light falls upon a transparent isotropic substance, it is refracted along a single direction, and the refracted ray obeys the law of refraction. When, however, a ray of light falls upon the surface of any transparent crystal other than one belonging to the regular system, it is in general divided into *two refracted rays*, one of which obeys the law of refraction and is called the *ordinary ray*, while the other follows a law of refraction altogether different from that of isotropic bodies, and is called the *extraordinary* ray. This phenomenon is called *double refraction.* It is exhibited by many animal and vegetable substances, and by glass, glue, gelatine, and similar substances when under stress. Double refraction is very readily observed in Iceland spar (crystallized $CaCO_3$), in which it was first observed in 1669, by Erasmus Bartholinus.

Iceland spar belongs to the hexagonal system of crystals and splits readily in planes corresponding to the three faces of a rhombohedron. Two of the solid angles which lie diametrically opposite are bounded by three equal obtuse angles of 101° 53′, while each of the remaining six are bounded by one obtuse and two acute angles. A line *aa′* (Fig. 301), making equal angles with the three faces forming the obtuse solid angles, is called the *optic axis* of the crystal.

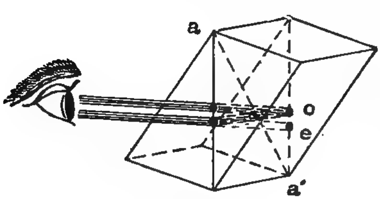


Fig. 301.

In the case of crystals which possess a principal axis of symmetry, a plane laid through this axis and including the normal to the surface, or any plane parallel to it is called a principal section. Consequently the plane which can be passed through the shorter diagonal of the rhomboidal surface and the axis of the rhombohedron is also a principal section. The rule for the double refraction in Iceland spar may be stated as

follows: A ray of light incident normally upon a rhomboidal surface of Iceland spar is separated into two rays, one of which vibrates at *right angles* to the principal section and is not deviated (ordinary ray), while the other vibrates *in* the principal section, and is deviated *in* the principal section, *away* from the end of the axis *aa'* (Fig. 301), in the face *toward which* it is going (extraordinary ray).

This may be illustrated by placing a rhomb of spar over a small black dot on white paper, and looking at the dot along the normal to the upper face of the crystal (Fig. 301). The eye will perceive two images, one in the continuation of the normal to the horizontal faces (ordinary image), and the other *in the principal section*, and *displaced from* the upper end of the principal axis (extraordinary image). On rotating the crystal over the dot the extraordinary image rotates about the ordinary image, but keeps its position relative to this image and the end of the axis *a* of the crystal.

510. Polarization by Double Refraction. If a ray of light be admitted through a small hole in a black card and a rhomb of Iceland spar be placed over it, the eye will perceive two rays emerging from the upper surface. These, tested either with a polarizing mirror, or a pile of plates, will show the following peculiarities:

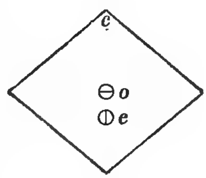


FIG. 302.

(*a*) The ordinary ray *o* (Fig. 302) will be found to be polarized *in the principal section of the crystal, i.e.* it swings at *right angles to the principal section*, while the extraordinary ray is polarized at right angles to the principal section, or its vibrations are *in the principal section.* Fig. 302 represents the front face of the rhomb in Fig. 301 and *c* denotes the end of the axis in the upper surface, designated by *a* in Fig. 301.

(*b*) If a second rhomb of equal thickness be similarly placed upon the first, that is, with the end of the optic axis in the upper face at *c* (Fig. 303) in each case, the same two images *o* and *e* (Fig. 303) appear as before, with similar

polarization, but the separation of the two rays is now twice as great as before. This is easily explained since the two rays emerge from the crystal as parallel rays and enter the second crystal in the same relative position in which they left the first. Hence the ordinary ray traverses the second crystal as ordinary and the extraordinary as extraordinary. The ordinary ray penetrates the crystal normally and hence suffers no deviation, while the extraordinary ray suffers the same deviation in the second as in the first; and since the plates are of equal thickness, the separation of the rays is twice as great for a rhomb of double the thickness.

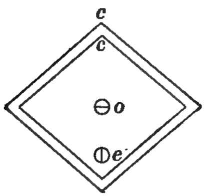


Fig. 303.

(*c*) Next, let the upper rhomb be rotated clockwise upon the lower, through an angle of about 30°. There will now appear *four* images instead of *two*, in the positions shown in Fig. 304. Upon examination it will be seen that the ordinary ray O in the first rhomb has been split into two by the second, producing an ordinary O_o and an extraordinary E_o, while the extraordinary in the first rhomb is likewise doubly refracted in the second, producing an ordinary O_e and an extraordinary E_e.

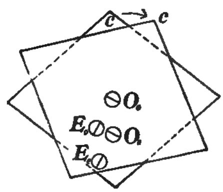


Fig. 304.

The intensities of the four rays increase and diminish in pairs. Thus, the two rays O_e and E_o are at first faint and gradually increase, while the pair O_o and E_e diminish in brightness, and when the crystal has been rotated through 90°, vanish entirely, leaving O_e and E_o (Fig. 305). For a rotation of 45° all rays possess equal intensity. After passing 90° the pair O_o and E_e reappear, and increase in brightness, the other pair diminishing to zero at 180°. At this point the two remaining

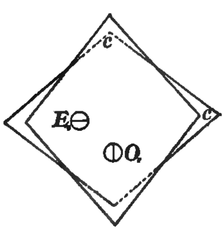


Fig. 305.

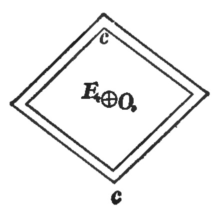


Fig. 306.

rays O_o and E_e coincide *for two rhombs of equal thickness*, and we have *a single image caused by two beams of light polarized at right angles to each other* (Fig. 306).

***511. Paths and Intensities of the Rays.** The paths of the rays through the two rhombs of equal thickness are shown in Fig. 307. The variation in the relative intensities of the four rays is readily understood from Fig. 308. Let JA and JB represent the directions of vibration of the ordinary and extraordinary ray in the first rhomb, and JD and JF the corresponding directions of vibration imposed upon the two rays of light upon entering the second rhomb, which has been rotated upon the first through an angle OJO'. Let JA and JB represent the intensities of the two rays on emerging from the first rhomb. By projecting JA and JB upon each of the new axes in turn, we have the four amplitudes of vibration,

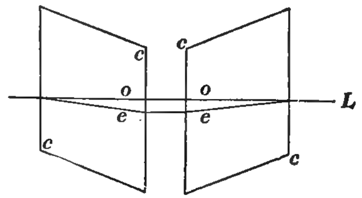


Fig. 307.

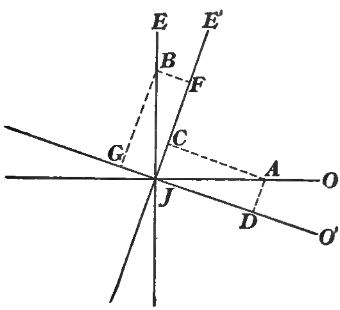


Fig. 308.

$$JD = O_o = JF = E_e$$

and $$JG = O_e = JC = E_o$$

whence the subsequent variations may be readily deduced.

***512. Indices of Refraction in Iceland Spar.** Since a ray of light in passing through a rhomb of Iceland spar is split up into two rays which are differently refracted, it follows that the crystal must have two indices of refraction. In determining the refractive indices of any uniaxial crystal it is convenient to employ the form of prism described in Article 452, and arrange the apparatus so that the incident light is normal to the first surface of the prism. In each case the direction of the optic axis in the prism is indicated by fine lines.

Three cases will be considered:

(*a*) *The optic axis of the crystal is parallel to the incident ray* (Fig. 309). In this case the ray traverses the crystal parallel to the optic axis and no double refraction results. The refractive index for sodium light is that for the ordinary ray, $\mu_o = 1.6585$.

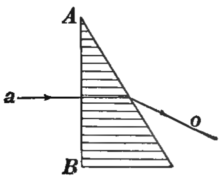


FIG. 309.

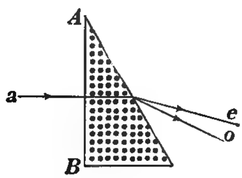


FIG. 310.

(*b*) *The optic axis is parallel to the refracting edge of the prism* (Fig. 310). Here the plane of incidence is normal to the axis, and consequently two rays emerge from the second side of the prism. The two values for the refractive index are, for the ordinary ray,

$$\mu_o = 1.6585$$

for the extraordinary ray, $\mu_e = 1.4865$

Should the angle of incidence change in this case, the direction of the ray through the prism would be changed, but the light would at all times traverse the crystal at right angles to the axis, and the values of the indices for the two rays would remain constant.

(*c*) *The optic axis is normal to the refracting edge of prism and parallel to first face of prism and to plane of incidence* (Fig. 311). In this case, so long as the incident light is normal to the first face of the prism, the same result will be obtained as in case *b*. Should the angle of incidence vary, however, the direction of the light through the prism is no longer normal to the optic axis, and the index for the ordinary ray alone remains constant, 1.6585, while the extraordinary index varies from 1.4865 to 1.6585.

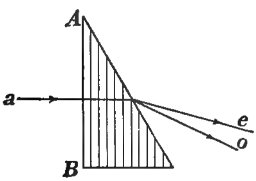


FIG. 311.

From this it is apparent that only in the second case would it be possible to determine the two indices by the method of minimum deviation, since in this case only does the light remain normal to the optic axis. In both the other cases the value of μ_e would vary between 1.4865 and 1.6585.

The results may be stated thus: A ray of light penetrating a crystal of Iceland spar is in general split into two rays. One, the ordinary ray, obeys the law of refraction and has always the same refractive index, $\mu_o = 1.6585$. For the extraordinary ray, the value of the refractive index varies between 1.6585 and 1.4865. It assumes the maximum value whenever the ray follows the optic axis, and the minimum value when it passes through the crystal at right angles to the optic axis.

***513. Wave Surfaces in Uniaxial Crystals.** Huygens explained the phenomena of double refraction in uniaxial crystals by an extension of the method adopted by him in the treatment of ordinary refraction. Since he had shown that the wave surface in an isotropic medium was a sphere, and since one of the rays in Iceland spar obeyed the laws of refraction in isotropic media, he assumed that for this ray the wave surface was a sphere. For the extraordinary ray he assumed the wave surface to be an ellipsoid of revolution about the optic axis, with its center at the point of incidence, and having for one of its axes the diameter of the sphere. Between

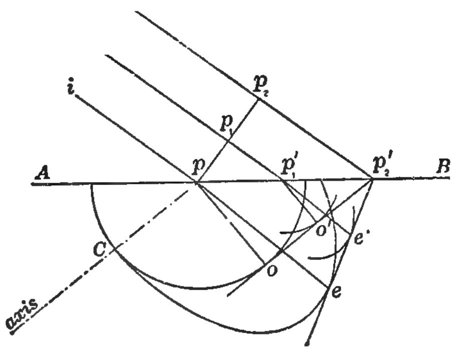


FIG. 312.

this axis and the second axis of the ellipsoid he assumed the same ratio to exist as existed between the velocities of the ordinary and extraordinary rays in the crystal. By means of these two surfaces the refracted waves may be found as in Article 486.

Thus, let pp_2 (Fig. 312) represent a plane wave front inci-

dent upon AB, the surface of a uniaxial crystal whose optic axis pC is assumed to lie in the plane of the paper. Then in time t, needed for the wave to travel the distance p_2p_2' *in air*, the spherical wave about p as center has reached o, and the spheroidal wave has spread to e. Then by the principle of Huygens, tangent planes $p_2'o'o$ and $p_2'e'e$, from p'_2 upon the sphere and spheroid respectively, mark the wave fronts of the ordinary and extraordinary waves in the crystal.

If the point of tangency to the sphere be o, then po is the ordinary ray, normal to the wave front. It lies in the plane of incidence and thus *obeys both laws of refraction.* If the tangent plane from p'_2 touch the spheroid at e, then pe is the extraordinary ray, which in general is not normal to the wave front and does not lie in the plane of incidence unless the optic axis is either *in* the plane of incidence or *normal* to it. In the figure the axis is assumed to lie in the plane of incidence, and hence the extraordinary ray also lies in that plane and so obeys one of the laws of refraction.

In one special case, however, *when the optic axis is normal* to the plane of incidence (Fig. 313), *the extraordinary ray obeys both laws of refraction;* that is, it is normal to the wave front and lies in the plane of incidence, and consequently its velocity in the crystal bears a constant ratio to the velocity in air, for all angles of incidence. This ratio is termed the extraordinary index of refraction μ_e.

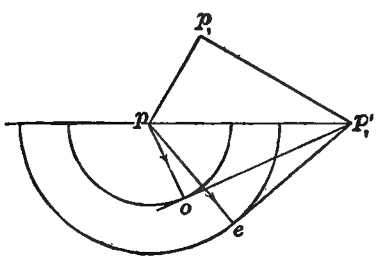


FIG. 313.

Thus (Fig. 313), the sections of the sphere and spheroid by the plane of incidence are both circles, and if the velocity in air be taken as unity, and the velocities of the ordinary and extraordinary rays as b and a respectively, then

$$\mu_o = \frac{p_1p_1'}{po} = \frac{1}{b} = \mu_o, \textit{ ordinary index} \qquad (591)$$

$$\mu_e = \frac{p_1p_1'}{pe} = \frac{1}{a} = \mu_e, \textit{ extraordinary index} \qquad (592)$$

This case corresponds to (*b*) of Article 512, in which it was shown that the values of both refractive indices were constant for all angles of incidence.

Crystals in which the extraordinary index μ_e is greater than the ordinary index μ_o are called *positive* crystals; those in which μ_e is less than μ_o are called *negative* crystals. In positive crystals, such as quartz, ice and zircon, the ellipsoid lies within the sphere; in negative crystals, as Iceland spar, tourmaline, beryl and sodic nitrate, the ellipsoid lies without the sphere.

CHAPTER LX

EXPERIMENTAL DEMONSTRATIONS

514. The Nicol Prism. We have seen that a beam of plane polarized light may be produced in any one of a number of ways, as by reflection from a polarizing mirror at an angle of 57°, by a pile of plates, by a crystal of tourmaline, or by double refraction through a crystal of Iceland spar. To all these methods there are more or less serious objections. In the polarizing mirror it is difficult to secure an intense beam of polarized light, since only about $\frac{1}{7}$ of the incident beam is reflected. In the pile of plates there is trouble from absorption by the plates, and diffusion of light from dust particles on the surfaces of the plates, thus causing stray light in the field. Tourmaline in plates of more than 2 mm thickness absorbs the ordinary ray completely, but has the disadvantage that the extraordinary ray which is transmitted is colored either green or red by the crystal, and also that tourmaline is not very transparent in plates of the required thickness.

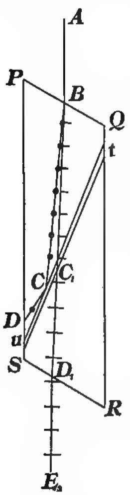


Fig. 314.

The most effective means of securing a strong beam of plane polarized light is by means of the Nicol prism. A clear crystal of Iceland spar is split out so that it is fully three times as long as it is broad. The end surfaces, which in nature make an angle of 72° with the edges of the side, are so cut as to make the angle *SPQ* (Fig. 314) 68°. Let the section *PQRS* represent a principal section of the rhomb. The prism is then sawed in two along a plane normal to the new end surfaces and to the plane of the principal section *PR*. The two new faces are then pol-

ished and cemented together with Canada balsam, which has a refractive index smaller than that of the spar for the ordinary ray, but larger than that for the extraordinary ray. If now a ray AB enter the rhomb parallel to its length the two rays are separated as usual in the spar, and the ordinary ray meets the Canada balsam at an angle slightly greater than the critical angle and is totally reflected through the side of the crystal and absorbed by its covering, which is painted dead black. The extraordinary ray passes into the Canada balsam as from a rarer to a denser medium and meets the second surface of the spar at an angle less than the critical angle from balsam to spar, and so is transmitted almost undiminished through the rhomb. As shown in the figure the vibrations in the extraordinary ray lie *in the principal section* through the shorter diagonal of the end surfaces of the rhomb. The Nicol thus transmits only those vibrations which are in the plane of its principal section, and quenches all vibrations at right angles to this plane.

515. Two Nicols. It is clear that if light emerging from one Nicol (polarizer) be passed through a second Nicol (analyzer), it will be transmitted if the principal sections of polarizer and analyzer be parallel, and will be totally extinguished if their planes be at right angles to each other. In any other position of the two Nicols there will be a portion transmitted and a portion absorbed. This is readily seen (Fig. 315).

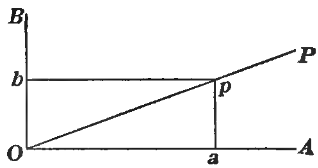


FIG. 315.

Thus, let a ray of plane polarized light, incident at O normally to the plane of the paper, have its vibrations parallel to OP, and let Op represent the amplitude of its vibration. Now, if OA represent the principal plane of the analyzer, it is clear that the component of vibration parallel to this plane is Oa, the transmitted portion; while the normal component Ob is absorbed by the analyzer.

If α or POA be the angle between the planes of the two Nicols, then Oa is $Op \cos \alpha$ and the intensity of the transmitted component is proportional to $\cos^2 \alpha$.

If a plate G (Fig. 316) of any *isotropic* substance be placed in a beam of plane polarized light between crossed Nicols, the field remains dark for any position of the plate rotating about the beam ST as an axis. The reason for this is found at once in the fact that an isotropic substance affects in no wise the directions of vibrations of light transmitted by it; it is not doubly refracting, and hence the plane polarized light from the polarizer is transmitted through it unchanged and is extinguished by the analyzer.

FIG. 316.

516. Doubly Refracting Substance in Parallel, Plane Polarized Light. If a thin plate of a doubly refracting crystal be interposed in a parallel beam of monochromatic light between crossed Nicols, there will be two positions of the plate in which the field will remain dark. These positions of the plate are those in which the two rectangular directions of vibrations in the crystal coincide with the planes of vibration in the two Nicols. In these positions, the light from the polarizer is transmitted unchanged by the crystal and quenched by the analyzer. In any other position of the plate as it is rotated about the beam as an axis, the field will light up. This is because the plane polarized vibrations from the polarizer are resolved by the plate into two component vibrations at right angles to each other, each of which will furnish a component parallel to the principal plane of the analyzer and so pass through, lighting up the field, while the other two components being at right angles to the plane of the analyzer, are quenched. If now white light be used instead of monochromatic light, the field will light up as before, on rotation of the crystal plate, but the light emerging from the analyzer will be colored and the color will depend upon the thickness of the crystal plate. If, on the other hand, the plate be fixed with its principal section at 45° to the plane of the polarizer and the analyzer be rotated, the color fades until white is reached at the position of coincidence of the plane of the analyzer with that of the plate; after passing this

position the color changes to the complementary hue and grows to a maximum saturation at 45° from the position for white. The complementary colors are therefore most pronounced when the planes of the Nicols are either parallel or crossed. This production of color from polarized light is due to interference. From the experiments of Fresnel and Arago, it was shown that two conditions were necessary for two beams of polarized light to interfere in the same way as in the case of ordinary light. *First, that the two beams of light shall be polarized in the same plane; second, that they shall have a common origin.*

Now, the two rays into which the light from the polarizer is split up by the crystal traverse the crystal plate with different velocities, and hence when the two components parallel to the plane of the analyzer are reunited, there will be a difference in phase, which for some color will amount to $\lambda/2$. The corresponding color will be absent, and the remaining light will be colored. The component from the extraordinary ray will be in advance of that from the ordinary ray, if the plate be from a *negative crystal;* it will be behind in phase if the plate be from a *positive crystal.*

517. Rings and Cross in Iceland Spar. If a thin parallel plate of Iceland spar or other uniaxial crystal be cut at right angles to the optic axis and interposed between crossed Nicols in a *convergent* beam of plane polarized light, there results a series of brilliantly colored concentric rings about a dark center and traversed by a dark cross, as shown in Fig. 317. The axis of the convergent beam should strike the plate normally, in which case it is transmitted along the axis of the crystal, suffers no double refraction, and is quenched by the analyzer, thus forming the dark center. Since the plate is cut normally to the optic axis, all planes normal to the surface of the plate and passing through the axis are *principal sections.* The vibrations of the ordinary ray are *normal* to these planes or *tangential* to the system of circles, while the vibrations of the extraordinary ray are *in* these planes and

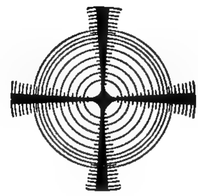


FIG. 317.

hence *radial* to the circles. Now, all rays of the convergent pencil, except the central one, will pierce the plate at an angle to the optic axis, and hence would ordinarily undergo double refraction; but in two of these planes, representing the planes of vibration of the polarizer and analyzer, one of the components is wanting. Thus if the polarizer transmit only vertical vibrations, then in the vertical diameter the vibration of the ordinary ray is zero, and in the horizontal diameter the vibration of the extraordinary ray is wanting, thus leaving only vertical vibrations in these two diameters which are quenched by the analyzer. In all other directions the rays traversing the plate furnish both extraordinary and ordinary rays which after resolution by the analyzer produce components vibrating in the same plane but differing in phase, owing to different speeds of transmission through the plate. Now the locus of all points in the plate producing any definite difference of phase, due to passage through a given thickness of the plate, will be a circle. Hence after the extinction of any wave length λ, there is left a circle of residual color through these points. Again, since a smaller thickness is needed to produce a difference of phase of $\lambda/2$ for violet than for red, it follows that the *violet* is extinguished *first* and the *red last* in each successive ring of color. The rings of residual color are therefore red on their inner and violet on their outer edges. Also, since the thickness of the plate traversed increases with the obliquity of the rays, the successive rings grow narrower toward the edges and by overlapping soon produce uniform illumination. If the analyzer be rotated through 90°, the black cross is replaced by a white one, and the rings are seen projected upon a white field instead of upon a dark background. The effect of superposing upon a uniaxial plate cut normally to the optic axis a second similar plate is the same as increasing the thickness of the first plate, *if they be from crystals of the same sign.* Hence a second negative plate upon the plate of Iceland spar would cause the rings to *contract.* A plate from a positive crystal would cause them to *dilate.* In this way the sign of a crystal may be determined by comparison with another crystal of known sign.

518. Double Refraction in Isotropic Media under Stress. If any isotropic substance be subjected to an unequal stress, the substance becomes doubly refracting and shows characteristic reactions when examined between crossed Nicols. Thus, if a piece of well-annealed glass, which shows no double refraction, be placed between the Nicols and subjected to slight pressure or tension, a characteristic colored pattern at once appears. If a narrow glass tube be placed between the crossed Nicols and then set in longitudinal vibration by stroking it with a moist cloth, the field will light up at each stroke of the cloth, thus showing the effect of the alternate compressions and dilations at the nodes of the sounding glass tube. Kundt showed that the glass tube behaved as a positive crystal (quartz) when dilating and as a negative crystal (Iceland spar) when contracting.

Mach has shown that viscous substances, like Canada balsam, warm rosin, hot glass, etc., may be made doubly refracting for a few moments by pressure or distortion. The effect passes away as soon as the molecules readjust themselves. Kerr has also shown that fluid as well as solid dielectrics become doubly refracting when subjected to electrical stress. When isotropic water freezes into ice, the direction of the optic axis is coincident with the direction of the *stress* due to gravity.

***519. Elliptic Polarization.** We have seen that if a beam of plane polarized light fall upon a thin plate of doubly refracting crystal, the original vibration is resolved into two vibrations at right angles to each other. These two sets of vibrations travel through the plate with different velocities, and consequently on emerging from the plate one vibration is behind the other in phase. Now, *within the region in which the two rays of light overlap*, the ether particles are simultaneously subjected to these two simple harmonic vibrations at right angles to each other, which have the same period, but which, in general, have different amplitudes and some definite difference in phase dependent upon the thickness of the plate. The motion resulting from compounding two such vibrations of the same period is an elliptic motion. Hence the light emerging

from the plate is said to be elliptically polarized *in the region where the two beams of light overlap.* This means that the paths of the ether particles transmitting the luminous disturbance are ellipses. Polarized light which has been reflected from a metal surface is also in general elliptically polarized. In the special case in which the two rectangular components have the same amplitude and a difference of phase of one quarter of a period, the resultant is circular motion. Circularly polarized light is produced by passing plane polarized light through a sheet of mica of such thickness that the retardation of one set of vibrations with respect to the other is *a quarter wave length.* Such a plate is called *a quarter wave plate.* Light which has been circularly polarized will appear equally bright for all positions of the analyzer. Elliptically polarized light shows maximum brightness when the plane of the analyzer is parallel to one axis of the ellipse, and minimum brightness when it is parallel to the other.

***520. Rotary Polarization.** If a plate of quartz cut normally to the optic axis be interposed between a pair of crossed Nicols, the darkened field is at once lighted up. If monochromatic light be used, the field can again be darkened by rotating the analyzer through a certain angle α. This shows that the light emerging from the quartz plate is still plane polarized, since it can be extinguished by the analyzer, but that the *plane of polarization has been rotated by the quartz plate through the angle α.* Experimental measurements have shown that the rotation is proportional to the *thickness of the quartz plate*, and *increases as the wave length decreases*, being about three times as great for violet light as for red. If, therefore, white light be used and the quartz plate be interposed between the crossed Nicols, the field will be colored, the tint depending upon the thickness of the plate; the color will remain no matter how the quartz plate be rotated about the beam as an axis. If the analyzer be rotated, the separate colors may be extinguished, one after the other, as the angle of rotation for the various colors is reached, but the field remains constantly lighted with the residual color. Some varieties of quartz rotate the plane of polarization *clockwise*

and are termed *right-handed*, others rotate the plane in the opposite direction and are called *left-handed* quartz. Besides quartz, many other crystals, as cinnabar, sodic chlorate, and the hyposulphates of lead, calcium and potassium possess this rotary power. Not only this, but even liquids and vapors have been found to possess rotary power, though in a much less degree. This seems the more remarkable, in view of the fact that *crystals when fused lose their rotary power.* Of the rotary active substances, perhaps the sugars are the most important, and the methods of testing and determining percentages of sugar in solution form one of the most important commercial applications of polarized light.

***521. Magneto-optical Rotation.** In 1845 Faraday discovered that isotropic substances, especially substances having high refractive power such as dense glass, were capable of rotating the plane of polarization of light when they were placed in a strong magnetic field. The rotation is *clockwise, to a person looking along the lines of induction, and in the direction of propagation of the light.* When the light is made to retrace its path by reflection, the direction of rotation is reversed, and so *the effect is increased proportionally to the number of times the light traverses the isotropic substance.* No effect is produced if the light pass at right angles to the lines of induction. The radical difference between this and other rotary phenomena is that in crystals or liquids the rotation produced by passing a beam through the substance is reduced to zero if the ray be made to retrace its path, while in the case of a magnetic field the rotational effect is doubled.

RADIATION

CHAPTER LXI

FUNDAMENTAL LAWS OF RADIATION

522. Introduction. Electrical and optical phenomena have been explained upon the assumption of certain disturbances in the ether. These disturbances may be of two kinds: *a*, static deformations or strains, causing magnetic and electrostatic phenomena, and *b*, dynamic disturbances, which are transmitted through free space with a speed of 3×10^{10} cm per second. As shown in the chapter on light (Art. 482), these dynamic disturbances are similar to wave motions in elastic bodies, and are therefore called ether waves. Just as in the case of sound, these disturbances spread radially from the source in all directions, and the resulting wave motion is always accompanied by a transfer of energy through space (Art. 105). This form of energy is therefore called *radiant energy*. Though, strictly speaking, sound waves are phenomena of radiation, the term is usually restricted to *ether radiations*. Only transverse ether waves are known at the present time (Art. 505).

Of the underlying causes of ether radiations we know but little. We shall see that all bodies, whether they be luminous or not, emit these radiations. The simplest assumption is, that in some manner, for example, by increased temperature or by electrical impulses, the motions of the molecules or of the electrons contained in the molecules are modified so as to throw the surrounding ether into oscillations, which, according to their mode of production, may be of widely different period and wave length.

Though ether waves of different wave length may necessitate the use of entirely different methods of observation and

measurement, they are all subject to the same general laws, most of which have already been mentioned in a somewhat different connection. In fact, *the only essential difference between ether waves is that of wave length.*

523. Methods of Observation. The following methods are those most frequently employed for the study of radiant energy.

(*a*) The most convenient method is direct *visual observation.* But, like the ear, the eye is limited in its range of sensibility and responds only to those vibrations whose wave lengths lie between 0.000812 mm and 0.000330 mm.

(*b*) For the detection of vibrations of wave length, shorter than those of the visible spectrum, the *photographic plate* may be used. This is most sensitive for the blue and the violet waves, and for those of still shorter wave length lying outside the visible spectrum. This invisible region is called the ultra-violet end of the spectrum, and by producing and photographing the spectra in vacuo, to avoid absorption by the air, Schumann succeeded in extending this end of the spectrum to a wave length of 0.0001 mm. By the application of certain dyes the photographic plate may also be made sensitive for the red end of the spectrum, but it cannot be used for waves lying below the visible spectrum.

(*c*) Short waves excite *fluorescence* in certain substances and the ultra-violet portion of the spectrum may be made visible by projecting it upon a fluorescent screen (Art. 543).

(*d*) The most important method for the investigation of ether radiations makes use of the *heating effect* produced when ether waves are absorbed by a black surface. Many different forms of instruments have been designed for this purpose. The oldest and best known is the thermopile (Art. 306). If this be connected to a sensitive galvanometer and placed in a spectrum, the deflection of the galvanometer indicates the amount of energy falling upon the instrument in the given position. By moving the thermopile through the whole length of the spectrum, the *energy curve* of the source of radiation may be plotted in terms of the wave length.

Since this method is independent of the wave length it is applicable for all parts of the spectrum. It is most useful for the investigation of the infra-red end of the spectrum, that is, the portion containing waves longer than the extreme red rays of the visible radiation.

(*e*) By *electrical means* ether waves may be produced whose lengths are very large in comparison with those of light. Since the production as well as the detection of these electric waves require the use of apparatus very different from those mentioned above, a special chapter will be devoted to their study (Chapter LXIV).

524. Radiation Spectrum. As has been pointed out in the discussion of different types of visible spectra (Arts. 495–500) different sources of radiant energy produce spectra of very different appearance. The same holds for all kinds of radiation spectra. The distribution of energy in a spectrum depends greatly upon the source. Thus while the maximum energy of the radiation from the sun lies in the greenish blue, the energy from an arc light or any other terrestrial source reaches its maximum in the infra-red.

The energy curve of the sun (Fig. 318) shows that only a small fraction of the total enegy emitted falls within the region of the visible spectrum. The sharp depressions in the curve mark the Fraunhofer lines, which occur as well in the invisible as in the visible parts of the spectrum.

Ether waves of widely different wave length have been investigated by the methods described in the last article. The shortest ether waves yet observed (Art. 523) are 0.0001 mm long. Within the past year Rubens and Wood[1] have detected and measured waves in the infra-red region of the enormous length of about 0.2 mm. These waves were detected in the spectrum radiated from an ordinary Welsbach burner. Also by the use of the quartz mercury lamp, Rubens and von Baeyer[2] have

[1] Rubens and Wood, *Ber. Ak. Wiss. Berlin*, Dec. 15, 1910; also *Phil. Mag.* 21, p. 249, 1911.

[2] Rubens and von Baeyer, *Phil. Mag.* 21, p. 689, 1911.

recently extended the spectrum to a wave length of 0.3 mm, so that at present the complete radiation spectrum from luminous bodies extends over about *twelve octaves*, or is *one thousand times as long as the visible spectrum*, which extends over but a little more than *one octave.*

The shortest electrical waves, thus far obtained, are 3 mm

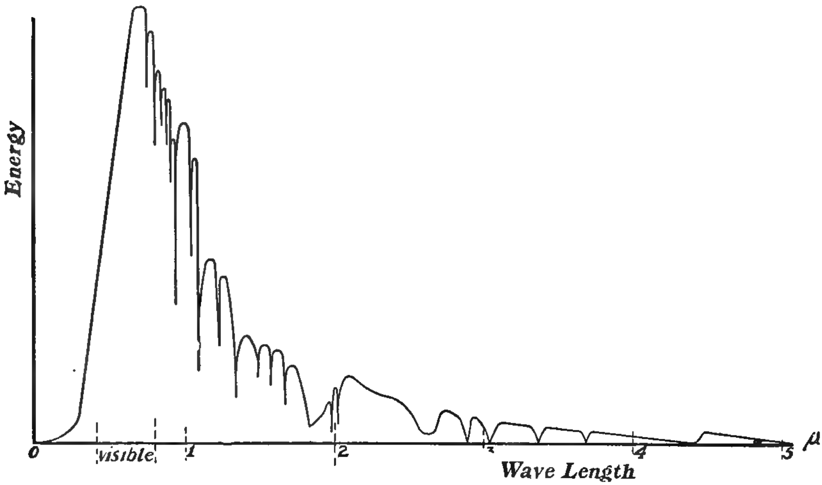


FIG. 318.

long. There is practically no limit to the length of electrical waves, since they may easily be made many miles in length. Thus we see that the region between the wave lengths of 0.3 mm and 3 mm is as yet unknown to us. This gap will surely be closed by future investigators.

The following table of wave lengths gives some idea of the range of the spectrum thus far explored, and of the relatively small part occupied by the visible spectrum. The wave lengths are expressed in millimeters as well as in μ. This unit is called the micron and is equal to 0.001 mm.

TABLE XIX

THE RADIATION SPECTRUM

	λ in mm	λ in μ
Shortest waves observed in vacuo	0.000100	0.1
Shortest waves observed in air	0.000185	0.185
Limit of visible light in the blue	0.000330	0.330
Blue hydrogen line	0.000486	0.486
Yellow sodium line	0.000589	0.589
Red hydrogen line	0.000656	0.656
Limit of visible light in the red	0.000812	0.812
Limit of line spectra observed	0.117	117
Longest waves from luminous bodies	0.3	300
Shortest electric waves	3.0	3000

525. Law of Inverse Squares. If *intensity of radiation* I be defined as the *quantity of radiant energy passing in unit time through unit area of a surface placed at right angles to the direction of propagation of energy*, the following law holds for all radiation, which proceeds from a point source: The *intensity of flux of radiation varies inversely as the square of the distance from the source* (Art. 108),

or
$$\frac{I}{I_1}=\frac{r_1^2}{r^2}$$

This law finds an important practical application in *photometry*, or that branch of physics which deals with the study and measurement of the luminous or subjective effects of radiation. Though the eye can form no exact estimate of the degree of intensity, it can determine with great accuracy equality of illumination upon two optically similar adjacent surfaces. If one of these surfaces be illuminated by one source of light and the other by a second source to be compared with the first, then equal illumination of the two surfaces may easily be obtained by varying the distances between the surfaces and the lights. The law of inverse squares then gives directly the ratio of the luminous intensities of the two sources in question.

526. Reflection and Refraction. If a heated iron ball be placed at the focus of a concave mirror, the radiation from the ball, after reflection from the surface of the mirror, forms a parallel beam. If another mirror be placed in the path of this beam at a distance of some 5 meters, the radiation will converge to the focus of the second mirror. A thermopile placed at this focus gives a deflection of the galvanometer, thus proving that the long waves obey the laws of reflection. The mirrors used in this experiment do not need to be highly polished (Art. 117), as in the corresponding experiment in light, since in this case we are dealing with much larger wave lengths.

That the law of refraction holds for radiation in general may be shown by placing a thin flask containing a solution of iodine in carbon disulphide in a parallel beam of light from an arc lamp. The solution absorbs nearly all the visible rays, but is transparent for the longer waves. These are focused a short distance behind the flask and a match may be lighted at the dark focus, showing a high concentration of energy at this place.

527. Interference, Diffraction and Polarization. Phenomena of interference are also found throughout the entire radiation spectrum. Huygens's principle may be applied in all cases of radiation. Moreover, all known ether waves may be polarized, either by means of doubly refracting substances or by apparatus, specially constructed for the purpose.

CHAPTER LXII

RADIATION AND TEMPERATURE

***528. Theory of Exchanges.** Any body at a temperature higher than that of its surroundings loses energy by radiation. The hand held a few centimeters from one face of a thermopile produces a deflection of the galvanometer on account of the passage of radiant energy from the hand to the thermopile. A non-luminous Bunsen flame affects the instrument through quite a distance. If a piece of ice be held in front of the thermopile, the galvanometer will be deflected in the opposite direction, showing that now the side of the instrument facing the ice is cooler than the other side. Since there are no "cold" rays, this reversal of the deflection can be explained only upon the assumption that the thermopile loses energy by radiation to the ice.

The theory of exchanges, first proposed by Prévost[1] in 1792, assumes that energy is radiated from every body and that radiant energy is also absorbed by every body. There is consequently a continuous interchange of energy between all bodies. If we confine ourselves to two bodies at different temperatures, the warmer radiates more energy than it absorbs, while the cooler body absorbs more than it radiates. The above experiments show, therefore, the differential effect between the emission and the absorption of radiant energy in the thermopile. We must assume that two bodies even at the same temperature continue to radiate and to absorb energy, but in this case just as much energy is absorbed by each body as is radiated by it.

529. Absorption and Emission of Radiant Energy. Of the total radiant energy falling upon any body, a part is reflected,

[1] Prévost, *Sur l'Equilibre du Feu*, Genève, 1792.

a part is absorbed and, in certain cases, a part is transmitted. The radiant energy absorbed by any body appears as heat in the body. The *absorptive power* of a body, or the *coefficient of absorption*, A, is that fraction of the total radiant energy falling upon the body which is transformed into heat. Thus, if I be the total energy, and I' be the energy absorbed, then

$$\frac{I'}{I} = A \tag{593}$$

The absorptive power differs with different substances and depends largely upon the character of the absorbing surface. In general dull black surfaces absorb much more energy than bright, polished surfaces which reflect a large portion of the energy.

In a manner entirely similar to equation (593) the *coefficient of reflection*, R, may be defined as that fraction of the total radiant energy falling upon a body which is reflected, or if I'' denote the energy reflected, then

$$\frac{I''}{I} = R \tag{594}$$

Both the absorptive and the reflective powers of a body vary with the wave length of the incident radiation, and with the angle of incidence. In the case of opaque bodies

$$R + A = 1 \tag{595}$$

In the case of transparent bodies it is clearly evident that the term transparent must be used advisedly, since there is no known substance which is transparent throughout the entire range of radiations of all possible wave length. All so-called transparent substances show strong absorption bands in some part of the radiation spectrum. Thus glass absorbs the violet and ultra-violet waves powerfully; it also absorbs the long waves in the ultra-red as may be shown by the slight effect upon a thermopile of a non-luminous Bunsen flame when shielded by a plate of glass; quartz, while very transparent to the ultra-violet and to most of the visible spectrum, shows strong absorption in the infra-red, becoming opaque for wave lengths between $7\,\mu$ and $24\,\mu$, while beyond this region it is

again quite transparent for the longest waves yet observed. A thin sheet of hard rubber is quite transparent to infra-red rays; thick black paper such as used for the protection of photographic plates against light waves transmits more than *thirty-three per cent* of those radiations whose wave length is 0.18 mm, and *seventy-nine per cent* of those of wave length 0.3 mm. For electric waves a two-inch plank or a brick wall are quite transparent. This property of showing well-defined absorption bands in definite parts of the spectrum is termed *selective absorption.*

All bodies emit radiant energy, the intensity of the radiation depending upon the temperature, the nature of the body and the condition of its surface. This intensity serves as a measure of the *emissive power* E of the body. Relative values of the emissive power of different bodies may be easily found by means of a Leslie's cube. This is a hollow cube whose sides are of different metals either polished or rough, or covered with lampblack. If the cube be filled with hot water, the surfaces are all at the same temperature. If, however, the cube be placed in front of a thermopile, the quantity of energy from the different faces of the cube will be very different. The dull black surface produces the largest deflection of the galvanometer, the rough metallic surfaces a much smaller deflection, while the polished metal surfaces have the least effect of all.

530. Kirchhoff's Law. From a study of the absorptive powers of glowing gases, Kirchhoff deduced a law connecting the radiation and absorption of bodies. This law may be stated as follows: *The ratio between the absorptive power and the emissive power is the same for all bodies at the same temperature, and the value of this ratio depends only on the temperature and the wave length.* A *black body* is defined as one which absorbs all the radiant energy which falls upon it. Hence it does not reflect or transmit any radiant energy. For such a body, therefore, *the absorptive power* $\frac{I'}{I}$ *is unity for all wave lengths and for all temperatures.* A small hole in a closed vessel is a close approach to a black body.

Now, according to Kirchhoff's law, the ratio $\frac{E}{A}$ for one body is the same as that for any other; hence

$$\frac{E}{A}=\frac{e}{1} \tag{596}$$

where e is the emissive power of the black body for the given wave length and temperature. From this it follows at once that the ratio $\frac{E}{A}$ for *any body* is equal to the emissive power of a *black body* for the same wave length and temperature. This means that if a vapor, as sodium vapor, at a given temperature emit yellow light more abundantly than other colors, it will also absorb that same color more abundantly, since

$$\frac{E}{A}=e=constant \tag{597}$$

531. Spectral Distribution of Energy. Since for every wave length the radiation emitted by any body depends upon the temperature only, the distribution of energy in the spectrum of the body is completely determined by the temperature of that body. The only body for which this distribution of energy in the spectrum may be derived from theoretical considerations is the black body. This fact gives the black body its important position among sources of radiant energy.

For other bodies the distribution of energy emitted at any temperature differs from the distribution of energy emitted by a black body at the same temperature. Further, since A is smaller than unity for all ordinary bodies, and since

$$E=Ae$$

it is evident that the energy radiated by any body at any temperature is always less than the energy radiated by a black body at the same temperature, provided the radiation be due to temperature alone.

532. Stefan's Law. Stefan's law [1] gives a quantitative measure of the total radiation emitted by a black body at different temperatures. Stefan's law states that *the total energy radiated*

[1] Stefan, *Wien. Ber.* 79, p. 391, 1879.

by a black body is directly proportional to the fourth power of the absolute temperature of the radiating body, or

$$E = C_1 T^4 \tag{598}$$

In the case of two black bodies, at absolute temperatures T_1 and T_2, Stefan's law takes the form

$$E = C_2(T_1^4 - T_2^4) \tag{599}$$

where C_2 is a new constant depending upon the area of the radiating surface as well as upon the temperature of the bodies which receive the radiation and which, in their turn, emit radiant energy proportional to the fourth power of their absolute temperatures.

Equation (599) may be written

$$E = C_2(T_1 - T_2)(T_1^3 + T_1^2 T_2 + T_1 T_2^2 + T_2^3) \tag{600}$$

If the difference of temperature $T_1 - T_2$ be small, we may place T_2 equal to T_1 in the second parenthesis and write

$$E = C_2 T_1^3 (T_1 - T_2) = C_3(T_1 - T_2) \tag{601}$$

where C_3 is a third constant.

For short time intervals the rate of emission may be assumed to be constant, in which case the content of equation (601) may be stated as follows: The rate of change of the temperature of a body is proportional to the difference of temperature between the body and its surroundings. This is known as Newton's law of cooling. It holds fairly well under the conditions stated, but must be considered as being at best only a rough approximation.

In all cases it should be borne in mind that Stefan's law (598) applies rigorously only for radiation from black bodies, and that it is liable to lead to considerable errors if applied without modification to other sources of radiant energy.

*** 533. Wien's Displacement Law.** As mentioned (Art. 531), the distribution of energy in the spectrum of a black body is a function of the temperature only, and for a given temperature the energy curve should have a perfectly definite form. The

law expressing such a relation was first stated by Wien,[1] and may be written

$$\lambda_m T = K \qquad (602)$$

where K is a constant, T denotes the absolute temperature and λ_m denotes the wave length corresponding to the maximum of the energy curve.

If λ be expressed in microns, or thousandths of a millimeter, the constant K for black bodies has the value 2900.

This law, if written in the form

$$\lambda_m = \frac{K}{T} \qquad (603)$$

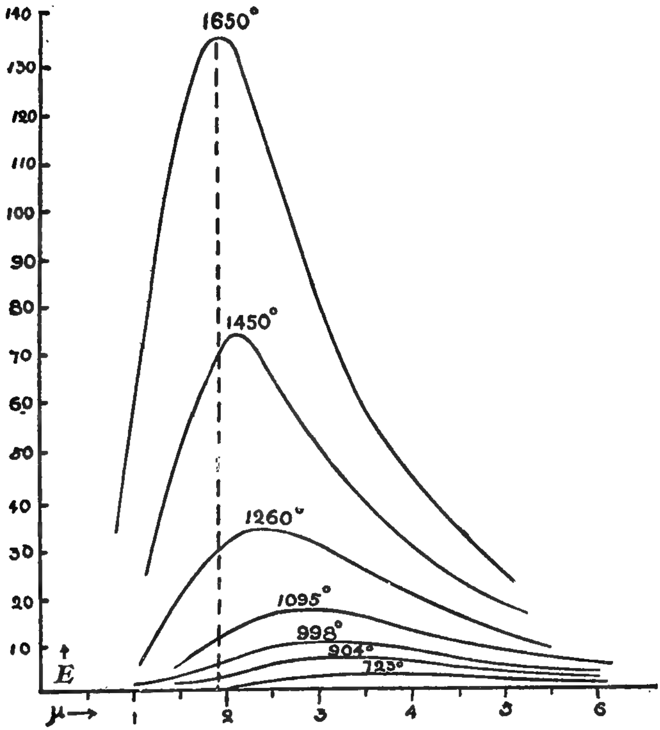


FIG. 319.

shows that *the maximum of the energy curve is displaced towards the shorter wave lengths as the temperature of the body rises.*

[1] Wien, *Ber. Ak. Wiss.* Berlin, 1893, p. 55.

For low temperatures (Fig. 319) all radiation lies in the infra red or the dark part of the spectrum. As the temperature rises, the energy curve extends into the visible spectrum and the body begins to emit light.

A black body becomes dull red at 525°C, yellow at 1000°C and white at about 1200°C. Other bodies begin to glow at about the same temperature as a black body, but the temperature will be higher the smaller the absorption of the body for the red.

*** 534. Wien's Second Law; Planck's Law.** Wien's displacement law does not give the form of the energy curve, that is, it does not express the relation between the radiation of any given definite wave length λ, and the corresponding temperature. In 1896 Wien[1] proposed a second law, expressing this relation, which may be written

$$E = C\lambda^{-5}e^{-\frac{c}{\lambda T}} \tag{604}$$

where C and c are constants, e the base of the natural logarithms and T the absolute temperature. This law holds in the region of the visible spectrum, but does not fit the experimental results obtained with longer waves. Planck[2] modified Wien's law and gave it a form which agrees well with the experimental results throughout the spectrum. It may be written thus

$$E = \frac{C\lambda^{-5}}{e^{\frac{c}{\lambda T}} - 1} \tag{605}$$

From a theoretical point of view this is the most important law in the theory of radiation since it gives the complete energy curve of a black body for any given temperature.

535. Temperature Measurement by Radiation. The above laws apply strictly to black bodies only, but they hold approximately also for bodies which are nearly black. Such laws may therefore be used for the determination of temperatures which are either too high for ordinary methods of thermometry, or for

[1] Wien, *Wied. Ann.* 58, p. 662, 1896.
[2] Planck, *Ann. d. Phys.* 4, p. 553, 1901.

the determination of temperatures of celestial bodies. This has given rise to a branch of thermometry called *optical pyrometry.* If the constants of equations (598, 602, 605), as found for black bodies, be used, the temperatures so calculated are called *black body temperatures.*

The black body temperature of the sun has been found to be about 6000°C, while that of the electric arc is about 3500°C. The latter result agrees well with that obtained by Violle, who in 1893 measured the temperature of the arc by dropping small tips of white-hot carbon broken from the positive into a known mass of water, and noting the resulting rise in the temperature of the water. The arc temperature was then calculated in accordance with the method of mixtures (Art. 173).

***536. Radiation Pressure.** As early as 1619 Kepler suggested that light exerts a pressure upon all bodies upon which it falls, and he attempted to explain by such pressure the fact that the tails of comets are always directed away from the sun. During the last century a number of unsuccessful attempts were made to measure this pressure. Maxwell, from his electromagnetic theory of light (Art. 405), had calculated this radiation pressure. He found that for a *perfectly black* surface it should be numerically equal to the energy contained in unit volume of the transmitting medium. For a *perfectly reflecting* surface the pressure should be twice this value. In 1900 Lebedew[1] in Russia, and in 1901 Nichols and Hull[2] in the United States, measured radiation pressure and obtained results, agreeing very well with the theoretical value.

Suppose a plane wave to fall perpendicularly upon a perfectly black surface of area A. Denote the amount of energy contained in unit volume of the medium by E and the velocity of radiation by v. Then the energy which reaches the surface and is absorbed by it in time t, is

$$W = EAvt \tag{606}$$

If now during the time t the black surface be moved away

[1] Lebedew, *Jour. d. russ. chem. Ges.* 32, p. 211, 1900

[2] Nichols and Hull, *Phys. Rev.* 13, p. 307, 1901.

from the source of radiant energy through a distance s, the amount of energy absorbed will be diminished by

$$W' = EAs \tag{607}$$

and this energy of radiation remains unabsorbed in the space As, in front of the displaced surface. But by the law of conservation of energy this decrease W' in the energy absorbed by the surface must be equal to the work done in displacing the surface, or

$$Fs = PAs = W' \tag{608}$$

where P denotes the pressure upon the surface. From this we find

$$P = \frac{W'}{As} = E \tag{609}$$

The radiation pressure upon a black surface, placed normally to the rays, is equal to the radiant energy per unit volume, falling upon the surface.

Accurate measurements have shown that at the surface of the earth the energy flux from the sun is 1.95 calories per square centimeter per minute. This quantity is called the *solar constant* k, where

$$k = 1.95 \frac{\text{calories}}{\text{min cm}^2} = 13.65 \times 10^5 \frac{\text{ergs}}{\text{sec cm}^2} \tag{610}$$

Now, since the velocity of radiation is 3×10^{10} cm per sec, the pressure upon a black surface on the earth, due to the radiation from the sun, is

$$P = \frac{13.65 \times 10^5}{3 \times 10^{10}} = 4.55 \times 10^{-5} \frac{\text{dynes}}{\text{cm}^2} \tag{611}$$

The energy received from the sun is very large. Thus from (610) it is seen that the time rate of energy falling upon a square meter is 1365 watts or 1.8 horse power.

Since force due to pressure is proportional to the area of the receiving surface, or to the *square* of the linear dimensions of the body, and since gravitational attraction is proportional to the mass, or to the *cube* of the linear dimensions of the body, we see that, as the body decreases in size, the gravitational

attraction decreases more rapidly than the force due to radiation pressure. For sufficiently small particles, such as are known to exist in the tails of comets, the force due to the radiation pressure from the sun will become larger than the attraction due to gravitation, and the particles will be driven away from the sun instead of being attracted towards it.

CHAPTER LXIII

COLOR

537. Color Sensation. Color is merely a sensation, and has no physical existence. The phenomenon of color sensation is chiefly of physiological origin and interest. The main physical difference to be noted between ether waves is that of wave length. Color sensation is the response of the nerves of the retina to the stimulus of a definite number of ether waves per second. The sensation of color as reported by the eye is much simpler in nature than the sensation of tone as reported by the ear. Thus the ear, on receiving a complex musical sound, is able, not only to receive the sound as a whole, but also to analyze it and to recognize its individual components. The eye, however, can perceive only the general or composite effect of a mixture of the spectral colors as white light and is unable to analyze any color mixture into its specific components.

As already noted (Art. 464), any method of recombining the spectral colors emerging from a prism produces upon the eye the effect of white light. This same result may be obtained by combining in proper proportion the three so-called "primary colors," red, green and violet, or by combining any pair of colors, obtained by mixing these three colors in any order, so long as the proper proportion is maintained. Two colors, which when added together produce the effect of white light, are said to be *complementary*. Thus yellow and blue, red and bluish green, greenish yellow and violet, are pairs of complementary colors.

Overstimulation of the nerves of the retina by any color tends to produce in the eye the sensation of its complementary color when white light is again admitted. Thus if a brightly lighted tile floor, showing a blue pattern upon a yellow ground, be viewed attentively for a few seconds, the eye on being

turned toward a grayish white wall will see a *yellow* pattern upon a *blue* ground. Such effects are termed *afterimages* or *subjective colors.*

538. Mixing of Colors. It is important to note that the effect of many of the spectral colors may be reproduced by the combination of certain fundamental or primary colors. The simplest method of showing such combination is by projecting different colored lights upon the same portion of a white screen. Thus red and green, when added together in the proper proportion, give a bright yellow; violet and green give a cyan blue. The yellow and blue thus obtained, if made to overlap, produce white and are therefore complementary colors. The three colors, *red*, *green* and *violet*, are often called *the three primary colors.*

It is possible to obtain all natural colors by the addition of the three primary colors in varying proportions. This has found a practical application in Ives's three-color method for color photography. It is to be clearly understood that in the cases just considered, colors have been produced by the *addition of colors.*

539. Mixing of Pigments. In sharp contrast to the method of producing color by mixing of colored lights is the method of producing color by mixing of pigments. If a thin glass cell containing a dilute solution of copper sulphate be placed in a beam of white light, the screen is covered with bright blue light. If instead of the cell containing copper sulphate, a second cell, containing a solution of normal potassium chromate be used, the screen is colored a bright yellow. If now the two cells be placed together in the beam, the screen appears green. This effect is explained by the fact that the copper sulphate solution absorbs from the white light the red, orange and yellow components, and transmits violet, indigo, blue and green; while the potassium chromate solution absorbs the violet, indigo and blue and transmits all the other colors. Consequently, when the two cells are combined, the green is the only color transmitted by both cells, and hence the screen

is colored green. In this case we have obtained color by the *subtraction* of certain colors from white light; and the intensity of the residual green light is less than that of the light transmitted by either cell.

In the case of colored paints or pigments, the light falling upon the pigment is subjected to a similar process of subtraction of color. The light penetrates the pigment a slight distance and is then reflected. In the light reflected from a mixture of pigments, all those colors are absent which are absorbed by any of the components of the mixture. Thus blue paint absorbs all spectral colors except blue and some green; yellow paint absorbs all except the yellow and some green. In the mixture of the two, the blue rays are absorbed by the particles of the yellow paint, and the yellow rays are absorbed by the blue paint. The mixture therefore appears green, and the intensity of the reflected light is smaller than it would be from either component alone.

540. Color of Natural Objects. The color of natural objects is, as a rule, due to the selective absorption of certain colors of the incident light and to the diffuse reflection of the unabsorbed colors. Since the incident light penetrates some distance into the medium before reflection occurs, ordinary bodies have the same color by reflected light as by transmitted light. The reflected color is complementary to that of the absorbed light. Thus it is seen that the colors of natural objects are due to a *process of subtraction*, or suppression of certain colors of the incident light.

Of course the spectral colors, reflected from a given body, must be present in the incident light, if the body is to appear as it does in daylight. If the illumination be such that none of the incident rays can be reflected by the body, the body will appear *black*. Thus a blue ribbon appears perfectly black when illuminated by yellow sodium light, since in this light there are no blue rays to be reflected. In short, two conditions are necessary for a body to appear of a definite color: First, the body must receive light containing that color; second, it must absorb all colors except that color.

Usually body colors are a mixture of rays from different parts of the spectrum, but they give the impression of a single color, since the eye is unable to distinguish the components of a color mixture. The following experiment brings out this fact most strikingly. Place a ruby glass plate and a plate stained with methylene green in a beam of white light. These colored plates remove from the spectrum the characteristic green of plants, leaving the light a dull gray, a mixture of light from both ends of the spectrum. If now a plant be illuminated by this light in a darkened room, *the leaves will appear blood-red.* This red color is reflected by the chlorophyl of the plant, but it is usually obscured by the much stronger reflection of green.

541. Surface Color. Some substances appear one color by reflected light and another color by transmitted light. Examples are seen in metals, certain solid aniline dyes, and heavy lubricating oils (Art. 468). Thin films of gold transmit green light; films of silver transmit blue light. In these cases the light seems to be reflected from the very surface of the bodies, instead of penetrating to a certain depth into the body before being reflected. These surface colors are said to be due to *metallic reflection*, because they are characteristic of all surfaces having a metallic appearance.

***542. Reststrahlen.** For very long waves the methods of obtaining spectra by diffraction (Art. 494) cannot be applied, since the grating spectra overlap in this region, so that separate rays cannot be isolated and studied. The longest waves observed and measured by this method were only about 0.01 mm long. Rubens and E. F. Nichols[1] were the first to apply the principle of selective reflection for the isolation of radiations of long wave lengths. Their method was based upon the following observation. After light proceeding from any source whatever has been reflected repeatedly from surfaces of the same substance there finally remain a few quite homogeneous waves sifted out by selective reflection from the original mixture of waves, and their properties may then be investigated by

[1] Rubens and Nichols, *Phys. Rev.* 4, p. 314, 1897.

ordinary methods. These "residual rays" are generally called by their German name "Reststrahlen."

By this method Rubens and his colaborators were able to extend our knowledge of the infra-red spectrum far beyond the limits previously attained. The following table gives the wave lengths of some of these isolated lines, observed in the farthest infra-red region.[1]

TABLE XX

LINE SPECTRA OF LONG WAVE LENGTHS

SUBSTANCE	WAVE LENGTH IN MM		WAVE LENGTH IN μ	
Rocksalt	0.0169	0.0536	46.9	53.6
Sylvine	0.0620	0.0703	62.0	70.3
Potassium bromide	0.0756	0.0865	75.6	86.5
Potassium iodide	0.0967		96.7	
Calcite [2]	0.093	0.116	93.0	116.1

*** 543. Fluorescence and Phosphorescence.** Certain substances have the property of absorbing radiant energy of definite wave length and of emitting this energy as waves of entirely different lengths characteristic of the substance. If the body cease to emit light when the illumination is cut off, the body is said to be fluorescent. Fluorescence is not a rare phenomenon. Thus fluorescein and uranium glass show light green fluorescence, sulphate of quinine and kerosene a light blue fluorescence, while chlorophyl fluoresces a deep red.

The light emitted by fluorescent substances is always of greater wave length than that which is absorbed to produce the fluorescence. The fluorescent spectrum usually consists of a large number of overlapping bands. Since fluorescence is produced by light of short wave length, it appears most brilliantly in violet or ultra-violet light. Cathode rays and Roentgen rays also produce strong fluorescence.

[1] Rubens und Hollnagel, *Verh. d.d. phys. Ges.* 12, p. 83, 1910.
[2] Rubens, *Verh. d.d. phys. Ges.* 13, p. 102, 1911.

Phosphorescence is a similar phenomenon and consists in a characteristic glow of certain substances, such as the sulphides of barium and calcium when strongly illuminated. In these substances, however, the glow continues after the illumination has ceased. Balmain's paint, after having been exposed to an intense light, will continue to shine brightly in the dark for many hours.

A characteristic property of phosphorescence is that it is at first intensified for a very short time by red and infra-red rays, but that it is soon extinguished by the same rays. If a spectrum be projected upon a phosphorescent screen, a black strip on a bright background will soon appear at the portion of the screen occupied by the red and infra-red rays. If the spectrum be that of the sun, the Fraunhofer lines in this region will be observed as bright lines on a dark background, since the presence of these lines prevents the extinction of the phosphorescence.

CHAPTER LXIV

ELECTRIC WAVES

544. Electrical Resonance. It has been known for many years (Art. 402) that the discharges of a condenser are oscillatory in character, having a frequency given by equation (436)

$$n = \frac{1}{2\pi}\sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}$$

Such discharges must produce periodic changes in the ether strain about the condenser, resulting in rapid alternations in the magnetic induction. These alternations in their turn produce rapidly alternating electromagnetic induction in any neighboring circuit which will be especially strong if this circuit chance to have a natural vibration period of the same frequency, or, as we may say, if it be *tuned* to the same frequency. In this case resonance effects will be observed very similar to those studied in the section on sound (Art. 132).

Lodge[1] showed electrical resonance by the following experiment. Two Leyden jars of the same capacity are furnished

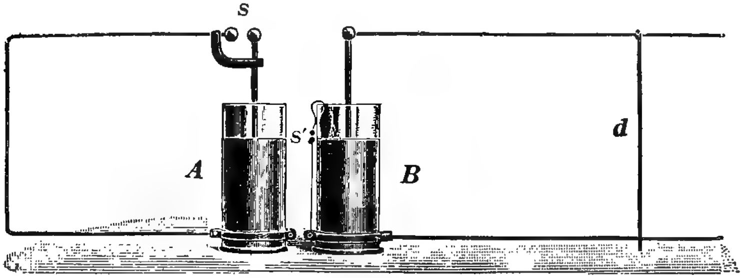


Fig. 320.

with discharging loops of nearly the same dimensions. The first circuit A (Fig. 320) contains a spark gap s, the second

[1] Lodge, *Nature*, 41, p. 368, 1890.

circuit B a spark gap of one or two millimeters at s' between the inner and outer coatings of the jar. The resistance and self-inductance of the circuit B may be varied by shifting the sliding wire d. By this means the frequency of oscillation of this circuit may be varied. If now the terminals of the first jar be connected to an induction coil, discharges take place across s, and, since the resistance in the circuit is so small as to render $\frac{R^2}{4L^2}$ very small in comparison with $\frac{1}{LC}$ in equation (436), these discharges are oscillatory.

By carefully adjusting the position of the sliding wire d in circuit B, this circuit may be tuned to exact resonance with the first. Forced vibrations are set up in circuit B, and sparks appear at the spark gap s' at discharge in circuit A. The two circuits may be thrown out of tune by moving the sliding wire from the position necessary for resonance. It is important to see just what happens in this experiment. The discharge in circuit A sets up vigorous oscillations in A. These oscillations are comparable to those of a tuning fork which has been struck by a soft hammer. These vibrations are quickly damped out, but, while they last, are sufficient to set up forced vibrations in a second tuning fork, represented by circuit B. In case resonance has been established between A and B, the vibrations in B become quite vigorous, producing sparks at s'. Each discharge in s corresponds to a new blow of the soft hammer, causing renewed sparks in s'.

A thick board or piece of glass may be placed between the two circuits without in the least affecting the experiment; but a sheet of thin metal suppresses the discharge in s' entirely. The electromagnetic action, therefore, *passes unhindered through dielectrics, but is completely stopped by conductors.*

From equation (437)

$$n=\frac{1}{2\pi}\sqrt{\frac{1}{LC}}$$

it is clear that, for a given frequency, if the capacity be small, the self-inductance must be correspondingly increased.

It is also clear that two circuits may have the same period without having the same capacity. If we use a coil of a great number of turns, we may dispense with a condenser entirely, since every coil has a small capacity. Such an arrangement was first used by Tesla, and is called a *Tesla coil.*

The discharging circuit consists of a Leyden jar or a plate glass condenser and a coil of a few turns of heavy wire. Inside this coil, called the primary from the similarity of the apparatus to the induction coil, is placed a coil consisting of a very large number of turns of fine wire, but having the same period of oscillation as the primary. This secondary coil responds strongly to the oscillations in the primary set up by each discharge of the condenser, and the potential difference between its terminals reaches enormous values. If the lower end of the secondary coil be connected to the earth, a brush discharge from six to ten inches long will be seen streaming out from the upper end.

In spite of its high potential, this apparently dangerous terminal may be touched without any ill effect upon the body. The frequency is so high that the electric disturbance in the medium does not find time to enter the human body. With lower frequencies, the same potential difference and current would cause instant death. The fact that quite large currents come into play in this experiment may easily be shown by interposing between the body and the secondary coil an incandescent lamp, which will be brightly lighted up.

545. Hertz's Experiments. Lodge's experiment was not designed to prove the existence of electric waves. In order to do this, it must be shown that the oscillatory disturbances in the ether travel with a finite velocity through space and obey the general laws of radiation. The first experimental proof of this was furnished by the celebrated researches of Hertz,[1] beginning in the year 1887.

If, as Maxwell had predicted, the discharges of a Leyden jar, or of any condenser, actually produce ether waves traveling

[1] Hertz, *Wied. Ann.* 31, p. 421, 1887.

through space with a velocity of 3×10^{10} centimeters per second, a simple calculation of the frequency of oscillation and an application of the wave formula

$$\lambda = \frac{v}{n}$$

will show that *the waves will be several miles long, even if small Leyden jars be used.* Hertz employed as discharging capacities small metallic spheres, short cylinders or plates, and was thus able in some of his experiments to reduce the wave length to about 60 cm.

In order to detect these waves, Hertz used a simple device, called *a spark-gap detector.* It consists of a rectangular or circular wire *dbc* (Fig. 321) containing a short spark gap *a*. The dimensions of this loop were so chosen that the detector was brought into resonance or *into tune* with the radiator *AB*. Then, when a discharge of sufficient intensity passed across *AB*, minute sparks also passed across the spark gap of the detector.

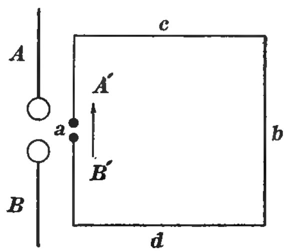


FIG. 321.

By his experiments Hertz proved that electric waves are reflected from plane and curved surfaces in accordance with the general laws of reflection. He refracted the waves through prisms made of dielectrics, such as resin, pitch or paraffine. He polarized them by passing them through coarse wire gratings and thus proved that, like all other ether radiations, they were of a transverse character.

By means of metallic mirrors Hertz produced a system of stationary waves by reflection (Art. 133), located nodes and antinodes, and measured the wave length of the waves in question. From the computed value for the frequency of the oscillator, and the measured value of the wave length, he showed that *the velocity of electric waves is the same as the velocity of light.* In short, he established the fact that electric waves exist, and that *the only difference between electric waves and other forms of ether radiation is one of wave length.*

546. Electric Radiators and Receivers. Most investigators after Hertz attempted to solve the following problems: To obtain shorter wave lengths, to increase the efficiency of the radiator, and to find more sensitive detectors for electric waves, in order to increase the accuracy of their experimental results.

By reducing the radiating system to minute dimensions, it has been made possible to produce electric waves of from 3 to 4 mm in length.

Many forms of radiators superior to that used by Hertz have been described. One of the most efficient is the Righi oscillator. It consists of three spark gaps instead of one. The central gap, which produces the effective waves, is immersed in paraffine oil.

Hertz's spark gap detector was not at all sensitive and was soon replaced by other receivers, for which the name *cymoscope* has been proposed. Of the many forms of receivers now in use we shall describe only the *coherer*. This was for some time considered the best receiver for electrical waves and is still much used for demonstration purposes.

The coherer consists of a tube containing metal filings between metallic electrodes. These filings form loose contacts and have a very high electrical resistance, so that a small difference of potential, such as that of a couple of dry cells, is insufficient to force an appreciable current through the coherer. But if the potential difference at the terminals of the coherer be raised to a sufficiently high value, the resistance is greatly decreased and the filings cling together. Thus if a coherer be placed in series with a cell and some current indicator, as a telephone or electric bell, no current will flow, owing to the high resistance of the coherer. But if an electric wave fall upon a wire connected to the earth through a coherer, the difference of potential between the wire and the earth may become high enough to break down the resistance of the coherer. In this case the cell is able to send a current through its circuit, actuate the telephone or electric bell and thus announce the arrival of the wave. After each signal the co-

herer must be tapped in order to separate the filings from each other and thus restore its high resistance.

For other forms of detectors, many of which are more sensitive and reliable than the coherer, the student is referred to textbooks on wireless telegraphy.

***547. Seibt's Experiments.** The following experiments, described by Seibt,[1] furnish a beautiful visible demonstration of electrical oscillations based upon the principle of resonance.

Two Leyden jars C_1 and C_2 (Fig. 322) are connected to an induction coil J and also to a discharging circuit, consisting of adjustable inductive resistances L_1 and L_2 and a spark gap F. By adjusting the sliding contacts K_1 and K_2 different lengths of the inductances are placed in the circuit. In this way the frequency of oscillation of the circuit may be varied within wide limits. To one of the terminals of the condenser a long vertical coil R of many turns is attached. This coil has a definite frequency and large differences of potential are produced by the electrical oscillations. Near this coil and parallel to it is placed a fine straight steel wire connected to the earth at its lower end, so that its potential is always kept at zero. If now the potential at any point of the coil reach a value sufficiently high, a brush discharge will be seen to pass between the coil and the wire at this place. The room must of course be darkened.

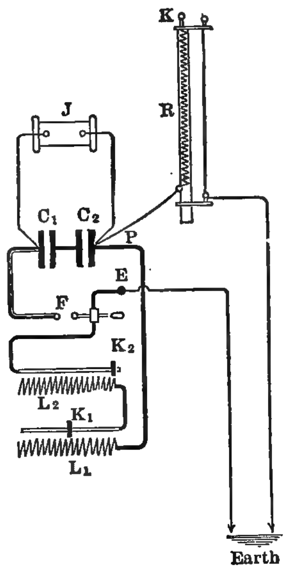


FIG. 322.

For the best effect the coil and the discharging circuit must be in tune, a condition easily obtained by adjusting the inductances in the circuit. Since the lower end of the coil is con-

[1] Seibt, *Phys. Zeitschr.* 4, p. 99, 1902.

nected to a large capacity, it forms an antinode or place of maximum freedom for the electrical surges. At the upper end *K* the coil is cut off, thus stopping the flow; hence there must be a node of electric displacement and consequently a high potential at this point. This high potential shows itself by a brush discharge to the wire. This is the simplest form of oscillation (Fig. 323) of the apparatus, and corresponds very closely to the vibrations in a closed organ pipe producing its fundamental tone (Art. 140, Fig. 73, *D*). For this lowest tone the condensers C_1 and C_2 must be connected in parallel, not in series, as shown in the figure. The vibration frequency may be increased by placing the condensers in series and adjusting the inductances. Thus overtones may be produced in the coil having one, two or three additional nodes and frequencies of 3, 5 or 7 times the frequency of the fundamental. A Geissler tube carried along the coil will brighten up at the nodes and remain dark at the antinodes. The difference of potential along the coil corresponds to the pressure variations in the organ pipe.

FIG. 323.

FIG. 324.

If the upper end of the coil be connected to the steel wire, this end will always be at zero potential and now becomes *an antinode*. With the fundamental vibration one node is seen at the center of the coil (Fig. 324).

Resonance effects appear more strikingly if two coils, having different numbers of wires and therefore different frequencies,

be connected at the same time to the discharging condenser. When the circuit is tuned to the frequency of either one, a strong brush discharge appears at the upper terminal of this coil, while the other remains dark (Fig. 325). If now a third coil, having the same frequency as one of these two coils, be placed at some distance from the apparatus, and its lower end connected to the earth, this coil will show resonance by a strong discharge at its upper end as soon as the discharging circuit is tuned to resonance with the coil having the same frequency; but it will not respond when the other coil is in oscillation.

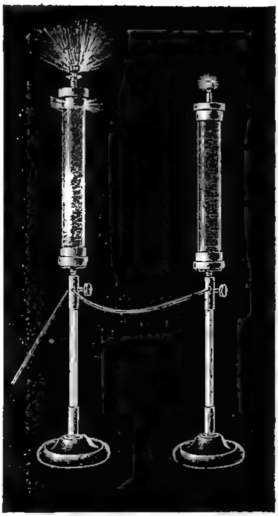


FIG. 325.

* **548. Wireless Telegraphy.** While Hertz's discovery contained all the possibilities of signaling through space by means of electric waves, the realization of practical wireless telegraphy required (*a*) the construction of a radiator which would allow a large portion of the energy of the oscillating system to be radiated into space, and (*b*) the use of a sensitive instrument for the detection of electric waves.

A nearly closed condenser circuit is a very poor radiator, since almost all the electrical energy of the discharge is transformed into heat in the metallic circuit. In general, short oscillating systems are poor radiators. The most effective radiators have an elongated form, because in this case the oscillations of the electromagnetic field surrounding the instrument are of large amplitude and pass out into space more freely than in the case of a short circuit. In practice, therefore, a part of the oscillator at the sending station always consists of a long wire, or system of wires, extending far above the ground. Such wires are called *antennae*. By this means large quantities of energy are radiated into space.

Of still greater importance is the use of a sensitive detector. Wireless telegraphy over long distances was made possible only after the invention of the coherer. With the detectors now in use messages may be sent across the ocean with the expenditure of much less energy than was needed for the earlier experiments over shorter distances. In most modern systems the sensitiveness is further increased by properly tuning the radiator and the receiver.

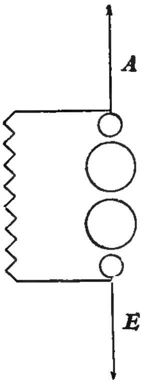


FIG. 326.

We shall describe only the principle underlying this new and important branch of applied electricity. At the sending station (Fig. 326) strong electric sparks are produced by an induction coil connected to the terminals of the radiator. In the figure Righi's form of spark gap is represented. One of the terminals of the spark gap is connected to the antenna A and the other to the earth E. At the receiving station (Fig. 327) a long antenna A is connected to the earth E through the detector c, represented in the figure as a coherer. When a wave is received by the antenna, the resistance of the coherer is lowered and allows the cell B to actuate the indicator in the circuit Bcr announcing the arrival of the wave.

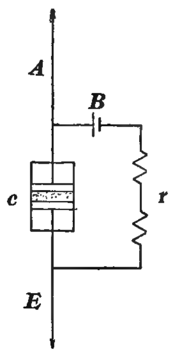


FIG. 327.

Communications with and between ships at sea form the most important practical application of wireless telegraphy.

*** 549. Wireless Telephony.** The electrical resistance of selenium varies greatly with the intensity of illumination falling upon it. This property may be used to detect a variation in the intensity of light. For this purpose two separate wires are wound parallel to each other and not more than a millimeter apart upon an insulating frame. Selenium is spread in the grooves between the wires and then heated to a little more than 100°C. Such an instrument is called a selenium cell, the free ends of the two wires serving as the terminals of the cell.

Since selenium is a poor conductor of electricity, the resistance between the terminals is very high, amounting to many thousand ohms in the usual forms. But the variation with different illumination is very great, the conductance of some cells increasing more than a hundred fold between darkness and daylight.

This peculiarity of the selenium cell is used in wireless telephony. At one station an arc light or other source of intense light is placed at the focus of a concave mirror which directs the beam of light to a similar mirror at the receiving station. A selenium cell is placed at the focus of the receiving mirror and is connected to a circuit, containing a battery and a telephone. Any variation in the intensity of light in the sending station will be *heard* through the telephone receiver at the other station.

***550. The Speaking Arc.** The arc light may be made to *speak* by the following simple arrangement. The arc A (Fig. 328) is connected to a source of direct current D, which for best results should be 200 volts or more. The resistances are so adjusted as to have at least 10 amperes flowing through the arc, which may be drawn out to a length of several centimeters. The circuit contains also a coil L of large self-inductance, which allows a steady current to pass, but offers a high resistance to rapidly changing currents (Art. 334), and is therefore called a choking coil. A shunt circuit, containing a telephone transmitter T, is attached to the terminals of the choking coil. The resistances of the coil and of the shunt circuit should be so chosen that the current through the transmitter which may be placed in a distant room will not exceed two amperes.

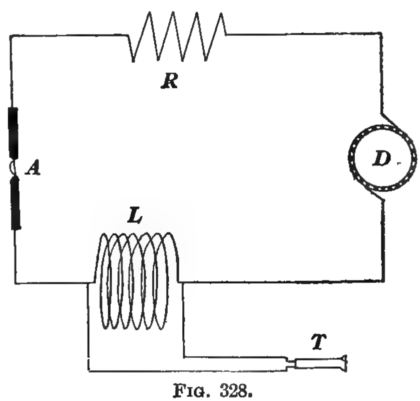


FIG. 328.

Speaking into the transmitter will produce rapid variations

of the current in the shunt circuit. These are forced by the choking coil through the arc and become superposed upon the steady current. The volume of the incandescent vapor of the arc is changed in unison with the variations of the current, giving rise to sound vibrations in the same manner as explained in the case of the singing arc (Art. 403). All sounds received by the telephone transmitter are reproduced by the arc with remarkable fidelity.

At the same time the intensity of light emitted by the arc varies in unison with the sound. It is evident that by the use of a speaking arc at the sending station the transmission of speech over large distances is made possible by means of wireless telephony.

***551. Index of Refraction for Ether Radiation.** Maxwell's theory (Art. 405), according to which all radiation is of an electromagnetic nature, was strongly supported by the experiments of Hertz. We have already seen that according to the electromagnetic theory the velocity of radiation v in any medium is

$$v = \frac{1}{\sqrt{c\mu}}$$

where c, the dielectric constant, and μ, the permeability, are measured in the same system of units. Again, it has been shown (Art. 486) that the relative index of refraction for two media is equal to the ratio of the velocities of light in the two media. Denoting the velocity of radiation in vacuo by V, and remembering that c is unity in vacuo and μ for transparent substances the same as in vacuo, we find for the absolute index of refraction n of any transparent medium,

$$n = \frac{V}{v} = \sqrt{c} \qquad (612)$$

Now, experiment has shown that for many substances the index of refraction, for light waves, is very different from the square root of the dielectric constant. But the dielectric constant is determined by means of very slow electric oscillations. If we wish to apply the above equation, the index of refraction must be calculated for very long waves. The best expression

for the relation between the index of refraction and wave length is the Ketteler-Helmholtz equation

$$n^2 = M_0 + \frac{M_1 \lambda_1^2}{\lambda^2 - \lambda_1^2} + \frac{M_2 \lambda_2^2}{\lambda^2 - \lambda_2^2} + \ . \ . \ . \qquad (613)$$

in which λ is the wave length for which n is to be found, λ_1, λ_2 . . . the wave lengths for absorption lines produced by the substance under investigation, and M_0, M_1 . . . constants, depending on the substance.

Applying this formula to very long waves the agreement between the electromagnetic theory and experimental determinations of the index of refraction is remarkably good. Thus for flint glass the index of refraction for sodium light is 1.62 and n^2 equals 2.62. For long waves n^2 increases to 6.7, while the dielectric constant has been found to lie between 6.7 and 8.

***552. Electron Theory of Radiation.** If we accept the electromagnetic theory of light, the question arises as to the manner in which the vibrations of short wave lengths, such as those of visible light, are produced. Since, according to this theory, radiation consists in a periodic disturbance of the electromagnetic condition of the ether, we must look for an explanation rather to an electrical disturbance in the source of light than to an elastic vibration of the atoms or molecules. Lorentz assumes that light is emitted by electric charges contained in the atoms of ponderable bodies. We may consider atoms as consisting of two parts; one, the larger portion, of the dimensions of an ion and charged positively, and a second part consisting of electrons or small negative charges which are in continuous vibratory motion about the positive center. The distribution of the electrons and their vibrations may be very complicated, but if we wish to explain the production of a single spectral line, we may assume that it is due to an electron vibrating with simple harmonic motion of a definite period. It is evident that such a motion would be least disturbed if the source of light were an incandescent gas. This, as we have seen (Art. 496), gives a line spectrum. It is not difficult to see that such a simple harmonic motion of an electron must pro-

duce corresponding electromagnetic disturbances in the ether about it and thus serve as a center of electromagnetic radiation.

This theory received remarkable experimental confirmation when Zeeman[1] found that line spectra are changed by placing the source of light in a strong magnetic field, and that the cause of this so-called *Zeeman effect* is a vibration of a negative charge of the magnitude of an electron, while no positive charges contribute to the radiation.

In accordance with this theory absorption of light and selective reflection are explained as resonance effects, producing sympathetic vibrations of the electrons contained in the absorbing substance. In short, the electron theory has substituted the negative electric charges for the vibrating material particles of the older theory of emission of light.

Since the electron theory explains also an electric current as being due to a motion of electrons *through* the conductor, we must expect a definite relation to exist between the electrical conductivity and the reflective and absorptive powers of a given substance. Experimental investigations have shown that for long waves, from 8 to 25 microns, the optical constants of metals may be calculated from their electrical conductivity and *vice versa*. The electromagnetic theory of light, as modified by the electron theory, has thus established a close connection between two groups of physical phenomena which at first sight would seem to be widely separated. Hence this theory is considered at the present time as the most satisfactory theory of radiation.

[1] Zeeman, *Phil. Mag.* 43, p. 226, 1897.

INDEX

Numbers refer to Articles

Printed in the United States of America.